

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**



## DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

(51) Classification internationale des brevets <sup>6</sup> : <b>C12N 15/00, 15/82, A01H 5/00</b>	<b>A1</b>	(11) Numéro de publication internationale: <b>WO 98/27202</b> (43) Date de publication internationale: <b>25 juin 1998 (25.06.98)</b>
(21) Numéro de la demande internationale: <b>PCT/FR97/02331</b> (22) Date de dépôt international: <b>17 décembre 1997 (17.12.97)</b> (30) Données relatives à la priorité: <b>96/16224</b> <b>17 décembre 1996 (17.12.96)</b> <b>FR</b> (71) Déposant ( <i>pour tous les Etats désignés sauf US</i> ): <b>BIOCEM S.A.</b> <b>[FR/FR]; Campus Universitaire des Cézeaux, 24, avenue des</b> <b>Landais, F-63170 Aubière (FR).</b> (72) Inventeurs; et (75) Inventeurs/Déposants ( <i>US seulement</i> ): <b>GRUBER, Véronique</b> <b>[FR/FR]; Résidence Sainte Madeleine, 44 C, avenue</b> <b>Jean-Jaurès, F-63400 Chamalières (FR). EXPOSITO,</b> <b>Jean-Yves [FR/FR]; Les Setives, F-38090 Bonnefamille</b> <b>(FR). RUGGIERO, Florence [FR/FR]; 26, rue d'Alsace,</b> <b>F-69100 Villeurbanne (FR). COMTE, Jeanne [FR/FR];</b> <b>"Le Brevent", 47, avenue Valioud, F-69110 Sainte Foy les</b> <b>Lyon (FR). GARRONE, Robert [FR/FR]; 26, rue Edgar</b> <b>Quinet, F-69500 Bron (FR). MEROT, Bertrand [FR/FR];</b> <b>3, La Coussedière, Moulet Marcenat, F-63530 Volvic</b> <b>(FR). BOURNAT, Philippe [FR/FR]; 48, rue Bonnabaud,</b> <b>F-63000 Clermont-Ferrand (FR).</b>	(74) Mandataire: <b>BREESE - MAJEROWICZ; 3, avenue de</b> <b>l'Opéra, F-75001 Paris (FR).</b> (81) Etats désignés: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY,</b> <b>CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH,</b> <b>HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,</b> <b>LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL,</b> <b>PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT,</b> <b>UA, UG, US, UZ, VN, YU, ZW, brevet ARIPO (GH, GM,</b> <b>KE, LS, MW, SD, SZ, UG, ZW), brevet eurasien (AM, AZ,</b> <b>BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE,</b> <b>CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL,</b> <b>PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN,</b> <b>ML, MR, NE, SN, TD, TG).</b> Publiée <i>Avec rapport de recherche internationale.</i> <i>Avant l'expiration du délai prévu pour la modification des</i> <i>revendications, sera republiée si de telles modifications sont</i> <i>reçues.</i>	

(54) Title: **RECOMBINANT COLLAGEN AND DERIVED PROTEINS PRODUCED BY PLANTS, METHODS FOR OBTAINING THEM AND USES**(54) Titre: **COLLAGENES RECOMBINANTS ET PROTEINES DERIVEES PRODUITS PAR LES PLANTES, LEURS PROCÉDES D'OBTENTION ET LEURS UTILISATIONS**

## (57) Abstract

The invention concerns the use of a recombinant nucleotide sequence containing a DNAC coding for one or several mammal collagen chains or derived proteins and elements enabling a plant cell to produce the collagen chain(s) or derived proteins, coded by said DNAC, particularly a transcription promoter and terminator identified by the transcription machinery of the plant cells, for transforming the plant cells so as to obtain from these cells, or plants obtained from them, the collagen chain(s) or derived proteins. The invention also concerns the resulting proteins and transformed plant material as well as their uses.

## (57) Abrégé

L'invention a pour objet l'utilisation d'une séquence nucléotidique recombinante contenant, d'une part, un ADNc codant pour une ou plusieurs chaînes de collagène de mammifères ou les protéines dérivées, et d'autre part, les éléments permettant à une cellule végétale de produire la ou les chaînes de collagène ou les protéines dérivées, codées par ledit ADNc, notamment un promoteur et un terminateur de transcription reconnus par la machinerie transcriptionnelle des cellules végétales, pour la transformation de cellules de plantes en vue de l'obtention, à partir de ces cellules, ou de plantes obtenues à partir de ces dernières, de la ou des chaînes de collagène ou des protéines dérivées; les protéines et le matériel végétal transformé ainsi obtenus ainsi que leurs utilisations.

Docket No.: **FG0219 US**Applicants: **Chang et al.**U.S. Serial No. **09/710,239**Filing Date: **10 November 2000**Title: **RECOMBINANT GELATINS**

# **UNIQUEMENT A TITRE D'INFORMATION**

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Bresili	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

COLLAGENES RECOMBINANTS ET PROTEINES  
DERIVEES PRODUITS PAR LES PLANTES, LEURS PROCEDES  
D'OBTENTION ET LEURS UTILISATIONS.

5 La présente invention a pour objet la production  
par les plantes de collagènes recombinants, notamment du  
collagène de type I homocatenaire [ $\alpha 1$  (I),] et autres  
dérivés polypeptidiques, ainsi que leurs utilisations.

On connaît dans l'art antérieur le brevet WO  
10 9603051 qui concerne la production de collagène dans le  
lait d'animaux transgéniques.

Le collagène est une protéine animale fibreuse  
extracellulaire, largement répandue dans les tissus  
animaux (récemment détectée également chez certains  
15 champignons). Certains organes en contiennent des  
quantités importantes : la peau (au niveau du derme), les  
tendons, les os. Il s'agit en fait d'un polymère dont les  
propriétés remarquables résultent à la fois des  
caractéristiques en triple hélice de certains domaines de  
20 sa molécule et de la régularité de ses assemblages supra-  
moléculaires. Elle est impliquée dans l'organisation de la  
matrice extracellulaire et regroupant une vingtaine de  
molécules différentes, appelées "types", et répertoriées  
en chiffres romains (actuellement de I à XIX) . Le domaine  
25 en triple hélice caractéristique est formé par  
l'enroulement de trois peptides, ou chaînes  $\alpha$ , organisés  
en hélice gauche et relevant d'une conformation unique,  
l'hélice  $\alpha$  du collagène. Cette conformation spécifique  
résulte de la répétition d'un triplet d'acides aminés,  
30 Gly-X-Y ou X est fréquemment représenté par la proline et  
Y par l'hydroxyproline. Ces aminoacides assurent la  
stabilité de ce type d'hélice. Dans une molécule de  
collagène, les trois chaînes  $\alpha$  (identifiées par un indice  
en chiffre arabe) disposées en une super hélice droite,  
35 peuvent être identiques ( $\alpha 1$  (ou  $\alpha 1$ )), de deux sortes

( $\alpha 1$  et  $\alpha 2$ ) ou toutes différentes ( $\alpha 1$ ,  $\alpha 2$ ,  $\alpha 3$ ). Les molécules de collagène comportent donc des domaines hélicoïdaux (ou domaines collagènes) et des domaines non hélicoïdaux. Elles s'associent en polymères homo ou hétérotypiques. Ainsi, les fibrilles de collagène qui constituent l'essentiel du derme sont-elles composées majoritairement du collagène de type I [ $\alpha 1$  (I),  $\alpha 2$  (I)] associé, à des collagènes de type III [ $\alpha 1$  (III)], de type V [ $\alpha 1$  (V),  $\alpha 2$  (V)] et recouvertes de collagènes de type XII ( $\alpha 1$  (XII),) et/ou XIV [ $\alpha 1$  (XIV),]. Des variantes peuvent exister : par exemple on connaît dans des tissus embryonnaires un collagène de type I homocatenaire [ $\alpha 1$  (I),]. Au cours de la biosynthèse du collagène, certains résidus prolyl et lysil sont hydroxylés, une addition de galactose, éventuellement complétée par un glucose peut se produire sur certains résidus hydroxyl et des N et O glycosylations classiques sont susceptibles d'intervenir sur les domaines non hélicoïdaux. La reconnaissance des trois chaînes formant une molécule et le début de leur assemblage sont sous le contrôle de l'extrémité C-terminale (C-propeptide). Le collagène de type I subit une coupure enzymatique de ses extrémités non hélicoïdales à la suite du signal peptide clivé précédemment, les N et C propeptides terminaux sont excisés au cours de la maturation du collagène laissant subsister des courtes extensions terminales non hélicoïdales (les télopeptides). Ce sont ces molécules clivées qui s'assemblent en polymères ordonnés (les fibrilles de collagène) et qui au cours du temps subissent une réticulation croisée via les résidus hydroxylysyl d'une molécule et les télopeptides d'une molécule adjacente. Les propriétés mécaniques et biologiques du collagène ont été exploitées depuis longtemps : une fois réticulé de manière irréversible (procédé de tannage) il fournit le cuir ; dénaturé par

chauffage, il donne naissance à la gélatine et à des colles. Mais ce n'est qu'au cours de la dernière décennie que le collagène a véritablement fourni des biomatériaux à usage pharmaceutique (compresses hémostatiques, éponges, 5 pansements notamment cicatrisants), médicaux (prothèses telles que valves cardiaques, tendons et ligaments, substituts cutanés, agents de comblement), odontologiques (implants gingivaux) et cosmétiques (additif, microconteneur de substances parfumées).

10 La meilleure connaissance ultérieure de cette protéine et des méthodes de purification ont conduit à la préparation de collagènes bovin et placentaire humain sous une forme prédéfinie : gel, éponge, poudre, fil, et microsphère par exemple. L'ingénierie des matrices 15 extracellulaires trouve également son application dans la constitution d'organoïdes renfermant des cellules transfectées pour des applications de thérapie génique par exemple. Le collagène majoritairement utilisé est le type I (en général associé à du type III) pour des raisons 20 d'abondance et de faible coût de purification et les sources principales en ont été bovine (peaux impropres à la tannerie) ovine (peau et intestins) et humaine (placenta). Cette dernière source ayant été exclusivement réservée à des applications pharmaceutiques ou médicales.

25 Alors que les propriétés mécaniques et biologiques très utiles du collagène sont reconnues sans ambiguïté, l'utilisation de cette protéine est remise en question en raison des risques éventuels de contamination par des agents infectieux non conventionnels. En effet, si 30 les risques posés par des contaminations bactériennes ou virales peuvent être parfaitement maîtrisés, il n'en est pas de même pour ceux associés aux agents de type prions. Ces agents infectieux qui semblent de nature protéique interviennent dans le développement d'encéphalopathies 35 dégénératives animales (tremblante du mouton, encéphalopathie spongiforme bovine) et humaines (syndrome de Creutzfeld-Jacob, syndrome de Gertstmann-Straussler,

kuru). Le délai de leur expression éventuelle est telle que les contrôles formels sont difficiles à réaliser. Ces risques ont déjà pratiquement bloqué toute commercialisation de collagène humain et les dispositions réglementaires édictées en ce qui concerne les collagènes animaux concernés compliquent les procédés de purification et augmentent leur coût.

Devant ces difficultés et face à une détérioration radicale de l'image du collagène de mammifère, une solution est la production de collagène recombinant que l'on pourrait purifier aisément dans un système non susceptible de provoquer des risques pathogènes pour l'homme et dont le coût industriel n'est pas prohibitif. Les inventeurs ont donc découvert et mis au point une production de collagène dans des espèces végétales. Par exemple, nous avons pu produire un collagène humain de type I. Sa molécule comporte une longue triple hélice ininterrompue et est très peu immunogène après purification. Les inventeurs ont par exemple fait exprimer la chaîne  $\alpha 1$  (I) afin d'obtenir des molécules homocaténaires  $\alpha 1$  (I), semblables à celles qui existent dans certains tissus, notamment embryonnaires.

Les cellules animales sont, a priori, plus adaptées à l'expression de gènes de mammifères. Leur utilisation pose cependant des problèmes de maturation des protéines. L'équipement enzymatique qui réalise la maturation post-traductionnelle est différent d'un tissu, d'un organe ou d'une espèce à l'autre. Par exemple, il a été rapporté que la maturation post-traductionnelle d'une protéine plasmatique peut être différente lorsqu'elle est obtenue à partir du sang humain ou bien lorsqu'elle est produite par une cellule recombinante comme les cellules d'ovaires de hamster chinois ou dans le lait d'un animal transgénique. Par ailleurs, les faibles niveaux d'expression obtenus avec les cellules des mammifères impliquent des cultures in vitro en très grands volumes à



des coûts élevés. La production de protéines recombinantes dans le lait des animaux transgéniques (souris, brebis et vaches) permet de diminuer les coûts de production et de surmonter les problèmes de niveau d'expression. Cependant,  
5 il reste des problèmes d'éthique et de contaminations virales et subvirales (prions).

Pour ces raisons, la transgénèse de gènes de mammifères dans une cellule végétale pourrait offrir une voie de production en grandes quantités de protéines  
10 recombinantes nouvelles, à un coût de production réduit et sans risque de contaminations virales ou subvirales. En 1983, plusieurs laboratoires ont découvert qu'il était possible de transférer un gène hétérologue dans le génome d'une cellule végétale et de régénérer des plantes  
15 transgéniques à partir de ces cellules génétiquement modifiées. Toutes les cellules de la plante possèdent alors le caractère génétiquement modifié qui est transmis à la descendance par fécondation sexuée.

Grâce à ces travaux, diverses équipes se sont  
20 intéressées à la production de protéines recombinantes de mammifères dans les cellules végétales ou dans des plantes transgéniques (Barta et al., 1986; Marx et al., 1982). L'un des premiers résultats vraiment significatif dans ce domaine a été la production d'anticorps dans les plantes  
25 de tabac transgéniques. Pour exprimer une protéine hétérologue dans la graine, lieu de stockage des protéines chez les végétaux, l'équipe de Vandekerckhove a fusionné la séquence codant pour la leu-enképhaline au gène codant pour l'albumine 2S d'*Arabidopsis thaliana*. Avec cette  
30 construction, des plantes d'*Arabidopsis* transgéniques ont été produites qui expriment la leu-enképhaline spécifiquement dans les graines à des niveaux d'expression de l'ordre de 0,1% des protéines totales. En 1990, le gène de la sérum albumine humaine a été transféré dans des  
35 cellules de tabac et de pomme de terre. Quelle que soit l'origine des peptides signaux (humaine ou végétale), des taux de sérum albumine humaine de l'ordre de 0,02% des

protéines totales ont été obtenus notamment dans les  
feuilles de pomme de terre. D'autres protéines  
recombinantes de mammifères ont été aussi produites dans  
les plantes: l'antigène de surface de l'hépatite B, les  
5 interférons, un anticorps de souris anti-*Streptococcus*  
*mutans*, agent de la carie, des fragments d'anticorps scFV  
anti-cellules cancéreuses, un anticorps anti-Herpès, la  
toxine du choléra et le facteur de croissance de  
l'épiderme humain (E.G.F). L'ensemble de ces recherches a  
10 permis de montrer que la production de protéines  
recombinantes de mammifères dans les cellules végétales  
est possible et que les mécanismes de la synthèse des  
protéines à partir des séquences d'ADN sont similaires  
chez les cellules animales et les cellules végétales. De  
15 nombreuses différences existent néanmoins entre les  
cellules végétales et animales, notamment au niveau de la  
maturation des glycanes polymannosidiques en glycanes  
complexes, ou encore au niveau des sites de clivage des  
peptides signaux, ne permettant pas ainsi de garantir  
20 l'obtention de protéines de mammifères actives ou  
suffisamment actives par transformation de cellules  
végétales.

Les inventeurs ont découvert que l'utilisation  
de cellules végétales transformées par une séquence  
25 nucléotidique recombinante appropriée, permet d'obtenir du  
collagène, notamment du collagène de type I homotrimérique  
[ $\alpha 1(I)$ ]<sub>3</sub>, recombinant.

Un autre but de la présente invention est de  
fournir les outils pour la mise en oeuvre d'un tel  
30 procédé, notamment de nouvelles séquences nucléotidiques  
recombinantes, des cellules de plantes transformées  
génétiquement, des plantes ou parties de plantes  
(notamment feuilles, tiges, fruits, semences ou graines,  
racines) transformées génétiquement, et des fragments de  
35 ces plantes ou parties de plantes transformées  
génétiquement.

L'invention a également pour but de fournir de nouveaux collagènes produits par les plantes, notamment du collagène de type I homotrimérique [ $\alpha 1(I)$ ],

L'invention a également pour but de fournir de nouvelles compositions protéiques susceptibles d'être utilisées dans le cadre de la mise en oeuvre ou la fourniture de compositions pharmaceutiques, médicales, odontologiques, cosmétiques, biotechnologiques ou industrielles.

10

#### DESCRIPTION DETAILLE DE L'INVENTION:

L'invention concerne :

- l'utilisation d'une séquence nucléotidique recombinante contenant, d'une part, un ADNc codant pour une ou plusieurs chaînes de collagène de mammifère  
15 notamment celle contenant à titre d'ADNc, celui de la chaîne du collagène  $\alpha 1$ , ou les protéines dérivées (on entend par protéine dérivée toute protéine présentant au moins 70 % d'homologie avec la protéine de référence, notamment au moins 80%, par exemple entre 85 et 100% d'homologie), et d'autre part, les éléments permettant à une cellule végétale de produire la ou les chaînes de collagène ou les protéines dérivées, codées par ledit ADNc, notamment un promoteur et un terminateur de  
25 transcription reconnus par la machinerie transcriptionnelle des cellules végétales, pour la transformation de cellules de plantes en vue de l'obtention, à partir de ces cellules, ou de plantes obtenues à partir de ces dernières, de la ou des chaînes de collagène ou des protéines dérivées, le cas échéant  
30 sous forme de triple hélice,

- une séquence nucléotidique recombinante, caractérisée en ce qu'elle contient d'une part la séquence codante pour une ou plusieurs chaînes de collagène de  
35 mammifère, notamment celle de la chaîne du collagène  $\alpha 1$ ,

ou les protéines dérivées et d'autre part, les éléments permettant à une cellule végétale de produire la ou les chaînes de collagène ou les protéines dérivées codées par ladite séquence, le cas échéant sous forme de triple  
5 hélice, notamment un promoteur et un terminateur de transcription reconnus par la machinerie transcriptionnelle des cellules végétales.

-un vecteur, notamment un plasmide, contenant une séquence nucléotidique selon l'invention insérée en un  
10 site non essentiel pour sa répllication.

-un hôte cellulaire, notamment toute bactérie telle que *Agrobacterium tumefaciens*, transformé par un vecteur selon l'invention,

-un procédé d'obtention d'une ou plusieurs  
15 chaînes de collagène ou des polypeptides dérivés, le cas échéant sous forme de triple hélice, caractérisé en ce qu'il comprend :

- la transformation de cellules végétales, notamment à l'aide d'un hôte cellulaire selon l'invention, lui-même transformé par un vecteur selon l'invention, de  
20 manière à intégrer dans le génome de ces cellules une séquence recombinante selon l'invention,

- le cas échéant, l'obtention de plantes transformées à partir des cellules transformées  
25 susmentionnées,

- la récupération de la ou des chaînes de collagène ou des polypeptides dérivés recombinants produits dans lesdites cellules ou plantes transformées susmentionnées, notamment par extraction, suivie, le cas  
30 échéant, par une purification,

-une plante, un extrait de plante ou une partie de plante, notamment feuilles et/ou fruits et/ou semences et/ou cellules de plantes, transformés génétiquement, caractérisée en ce qu'elle contient une (ou plusieurs)  
35 séquence(s) nucléotidique(s) recombinante(s) selon l'invention, intégrée(s) de façon stable dans leur génome, ces plantes étant choisies notamment parmi le colza, le

tabac, le maïs, le pois, la tomate, la carotte, le blé, l'orge, la pomme de terre, le soja, le tournesol, la laitue, le riz, la luzerne, et la betterave,

5 -une ou plusieurs chaînes de collagène ou protéine dérivée caractérisée en ce qu'elle est obtenue selon le procédé de l'invention,

-un collagène (notamment un collagène de type I, II, III, IV ou V) ou une protéine dérivée caractérisé en ce qu'il est obtenu selon le procédé de l'invention,

10 -un produit, notamment gélatine, caractérisé en ce qu'il est obtenu à partir des chaînes de collagène, du collagène ou de leurs protéines dérivées selon l'invention,

- une plante, un extrait de plante ou une partie  
15 de plantes, notamment feuilles et/ou fruits et/ou semences et/ou cellules de plantes, transformés génétiquement, caractérisés en ce qu'ils contiennent des chaînes de collagène, du collagène ou les protéines dérivées selon l'invention, ces plantes étant choisies notamment parmi le  
20 colza, le tabac, le maïs, le pois, la tomate, la carotte, le blé, l'orge, la pomme de terre, le soja, le tournesol, la laitue, le riz, la luzerne, et la betterave.

-l'utilisation de plantes, extraits de plantes ou parties de plantes selon l'invention, et/ou de  
25 protéines (chaînes de collagène, collagène ou protéines dérivées) selon l'invention, pour l'obtention de compositions pharmaceutiques, médicales, odontologiques, cosmétiques ou biotechnologiques,

-un Biomatériau et une composition  
30 pharmaceutique, médicale, odontologique, cosmétique ou biotechnologique. caractérisée en ce qu'elle comprend des plantes, extraits de plante, parties de plantes ou des chaînes de collagène, du collagène ou les protéines dérivées selon l'invention.

35 Une composition pharmaceutique selon l'invention comprend notamment toute composition selon l'invention constituant (ou entrant dans la fabrication de) une

composition permettant de prévenir ou de traiter toute pathologie liée à un dysfonctionnement du collagène, ainsi qu'une compresse pour cicatrisation de plaies, un pansement hémostatique, pansement anti-escarres, une  
5 poudre hémostatique ; ou un pansement pour brûlures.

Une composition médicale selon l'invention comprend notamment toute composition selon l'invention constituant (ou entrant dans la fabrication de) un dispositif de recouvrement cornéen, un film coagulant pour  
10 résection d'organes (foie en particulier), un matériau de comblement pour prévention d'adhérences et de fistules, une composition pour la confection de prothèses vasculaires et cardiaques (valvules), un conduit-guide pour régénération de nerfs, matériau de comblement osseux  
15 et intra corporel, un conteneur-relargueur de substances actives (hormones, facteurs de croissance, antibiotiques, anticancéreux, anti-inflammatoires par exemple), un dispositifs de compression (par exemple pour réduire l'incontinence urinaire), un substituts cutané, un fil  
20 chirurgical, un matériau injectable pour chirurgie esthétique (comblement de dépressions cutanées ou remodelage du visage par exemple).

Une composition odontologique selon l'invention comprend notamment toute composition selon l'invention constituant (ou entrant dans la fabrication de) un  
25 pansement gingival ou un matériel de comblement.

Une composition cosmétique selon l'invention comprend notamment toute composition selon l'invention constituant (ou entrant dans la fabrication de) un additif  
30 pour préparation (crèmes, pommades, fards, onguents).

Une composition médicale selon l'invention comprend notamment toute composition selon l'invention constituant (ou entrant dans la fabrication de) un système de : recouvrement de boîtes et flacons de culture  
35 cellulaire.

L'invention concerne également un produit, notamment la gélatine, caractérisé en ce qu'il est obtenu

à partir des chaînes de collagène, du collagène ou de leurs protéines dérivées selon l'invention.

En effet la gélatine est normalement préparée à partir du collagène par chauffage des tissus animaux (notamment des os et de la peau) dans l'eau puis séchage (Parkany M, 1984). Cette préparation détruit la structure secondaire du collagène et conduit donc à des changements importants de la solubilité et des propriétés mécaniques du produit. La gélatine est le principal constituant de la colle. Dans l'eau froide, elle gonfle et est insoluble. Dans l'eau chaude, elle se dissout pour donner une solution très visqueuse, qui se gélifie après refroidissement lorsque le taux de gélatine est supérieur à 1%. Pour la préparation de prothèses chirurgicales, des solutions contenant de 10 à 35 % de gélatine sont utilisées.

La propriété de la gélatine d'être soluble dans de l'eau chaude la rend utile pour de nombreuses applications alimentaires et industrielles. Elle est souvent préparée sous forme de poudre ou de fines feuilles. Suivant les propriétés mécaniques recherchées, du glycérol du sorbitol ou des agents de réticulations peuvent par exemple être ajoutés lors de la préparation des blocs de gélatines.

La gélatine est également utilisable dans le domaine bio-médical comme éponges par exemple.

Avantageusement, les séquences nucléotidiques recombinantes selon l'invention contiennent une (ou plusieurs) séquence(s) codant pour un peptide responsable de l'adressage des polypeptides recombinants dans un compartiment déterminé de la cellule végétale, notamment dans le réticulum endoplasmique ou dans les vacuoles, ou bien même à l'extérieur de la cellule, dans la paroi pectocellulosique ou dans l'espace extracellulaire aussi appelé apoplasme.

Parmi les terminateurs de transcription susceptibles d'être utilisés pour la transformation de

cellules de plantes dans le cadre de la présente invention, on peut citer le terminateur polyA 35S du virus de la mosaïque du chou-fleur (CaMV), ou le terminateur polyA NOS, qui correspond à la région en 3' non codante du  
5 gène de la nopaline synthase du plasmide Ti d'*Agrobacterium tumefaciens* souche à nopaline .

A ce titre, l'invention a pour objet toute séquence nucléotidique recombinante telle que décrite ci-dessus contenant en aval dudit ADNc ou de sa séquence  
10 dérivée, le terminateur polyA 35S du CaMV, ou le terminateur polyA NOS d'*Agrobacterium tumefaciens*.

Parmi les promoteurs de transcription susceptibles d'être utilisés pour la transformation de cellules de plantes dans le cadre de la présente  
15 invention, on peut citer :

- le promoteur 35S (P35S), ou avantageusement le promoteur constitutif double 35S (Pd35S) du CaMV, ces promoteurs permettant l'expression des polypeptides recombinants de l'invention dans l'ensemble de la plante  
20 obtenue à partir de cellules transformées selon l'invention, et sont décrits dans l'article de Kay et al., 1987,

- le promoteur PCRU du gène de la cruciférine de radis permettant l'expression des polypeptides  
25 recombinants de l'invention uniquement dans les semences (ou graines) de la plante obtenue à partir de cellules transformées selon l'invention, et décrit dans l'article de Depigny-This et al., 1992,

les promoteurs PGEA1 et PGEA6 correspondant à la  
30 région 5' non codante des gènes de la protéine de réserve de graines, GEA1 et GEA6, respectivement, d'*Arabidopsis thaliana* (Gaubier et al., 1993), et permettant une expression spécifique dans les graines,

- le promoteur chimérique super-promoteur PSP  
35 (Ni M et al., 1995), constitué de la fusion d'une triple répétition d'un élément activateur transcriptionnel du promoteur du gène de l'octopine synthase d'*Agrobacterium*



tumefaciens, d'un élément activateur transcriptionnel du promoteur du gène de mannopine synthase et du promoteur mannopine synthase d'*Agrobacterium tumefaciens*,

- le promoteur actine du riz suivi de l'intron actine de riz (PAR-IAR) contenu dans le plasmide pAct1-F4 décrit par McElroy et al. (1991),

- le promoteur HMGW (High Molecular Weight Glutenine) d'orge (Anderson O.D. et al, 1989),

- le promoteur du gène de  $\gamma$ zéine de maïs (Py zéine) contenu dans le plasmide py63 décrit dans Reina et al. (1990), et permettant l'expression dans l'albumen des semences de maïs.

A ce titre, l'invention a pour objet toute séquence nucléotidique recombinante telle que décrite ci-dessus, contenant en amont dudit ADNc ou de sa séquence dérivée, le promoteur constitutif double 35S (Pd35S) du CaMV, ou le promoteur PCRU du gène de la cruciférine de radis, ou les promoteurs PGEA1 ou PGEA6 d'*Arabidopsis thaliana*, ou le super-promoteur PSP d'*Agrobacterium tumefaciens*, ou le promoteur PAR-IAR du riz, le promoteur HMGW d'orge ou le promoteur  $\gamma$ zéine du maïs.

Les séquences codant pour un peptide d'adressage utilisées dans le cadre de la présente invention, peuvent être d'origine végétale, humaine ou animale.

Parmi les séquences codant pour un peptide d'adressage d'origine végétale, on peut citer:

- la séquence nucléotidique de 69 nucléotides (indiquée dans les exemples qui suivent) codant pour le prépeptide (peptide signal) de 23 acides aminés de la sporamine A chez la patate douce, ce peptide signal permettant l'entrée des polypeptides recombinants de l'invention dans le système de sécrétion des cellules végétales transformées selon l'invention (à savoir principalement dans le réticulum endoplasmique),

- la séquence nucléotidique de 42 nucléotides (indiquée dans les exemples qui suivent) codant pour le

propeptide N-terminal d'adressage vacuolaire de 14 acides aminés de la sporamine A chez la patate douce, permettant l'accumulation des polypeptides recombinants de l'invention dans les vacuoles des cellules végétales transformées selon l'invention,

- la séquence nucléotidique de 111 nucléotides (indiquée dans les exemples qui suivent) codant pour le prépropeptide de 37 acides aminés de la sporamine A constitué de la partie N-terminale vers la partie C-terminale des 23 acides aminés du peptide signal susmentionné suivis par les 14 acides aminés du propeptide susmentionné, ce prépropeptide permettant l'entrée de polypeptides recombinants de l'invention dans le système de sécrétion et leur accumulation dans les vacuoles des cellules végétales transformées selon l'invention,

les trois séquences susmentionnées étant décrites dans les articles de Murakami et al., 1986, et Matsuoka et al., 1991,

- le propeptide carboxyterminal de la lectine d'orge décrit notamment dans les articles de Schroeder et al., 1993, et de Bednarek et al., 1991,

- et le PRS (Pathogenesis Related Protein, Cornelissen et al. 1986) permettant la sécrétion.

On peut également citer, parmi les séquences codant pour un peptide d'adressage, celle codant pour les peptides KDEL, SEKDEL et HDEL et permettant un adressage dans le réticulum endoplasmique.

L'invention a également pour objet toute séquence nucléotidique recombinante telle que décrite ci-dessus, contenant une séquence codant pour tout ou partie d'un peptide d'adressage vacuolaire, notamment celui de la sporamine A de la patate douce, cette séquence codant pour un peptide d'adressage vacuolaire étant située, dans ladite séquence nucléotidique recombinante, entre la séquence codant pour un peptide signal et celle codant pour ledit ADNC ou sa séquence dérivée, de telle sorte que le premier acide aminé N-terminal du peptide d'adressage

vacuolaire soit lié au dernier acide aminé C-terminal du peptide signal, et que le dernier acide aminé C-terminal dudit peptide d'adressage soit lié au premier acide aminé N-terminal du polypeptide codé par ledit ADNc ou sa  
5 séquence dérivée, dans la protéine codée par ladite séquence nucléotidique recombinante.

L'invention a également pour objet toute séquence nucléotidique recombinante telle que décrite ci-dessus, contenant une séquence codant pour tout ou partie  
10 d'un peptide d'adressage vacuolaire, notamment celui de la lectine d'orge, cette séquence codant pour un peptide d'adressage vacuolaire étant située, dans ladite séquence nucléotidique recombinante, en aval de la séquence codant pour ledit ADNc ou sa séquence dérivée, de telle sorte que  
15 le premier acide aminé N-terminal du peptide d'adressage vacuolaire soit lié au dernier acide aminé C-terminal du polypeptide codé par ledit ADNc ou sa séquence dérivée, dans la protéine codée par ladite séquence nucléotidique recombinante.

20

#### DESCRIPTION DETAILLEE DES FIGURES

##### Fig 1:

Immunotransfert montrant un exemple de résultat positif issu du criblage de plants transformés par les  
25 constructions pBIOC706 (indiqué 706, plant n°45, pistes 2,3) et pBIOC707 (indiqué 707, plant n°2, pistes 4,5). Ces résultats sont à comparer avec ceux obtenus avec les plants témoins non transformés (pistes 6, 7). La piste 1 présente la migration électrophorétique du collagène I  
30 témoin: la bande correspondant à  $\alpha 2(I)$  migre au niveau de la bande 116 KDa des poids moléculaires standards, la bande  $\alpha 1(I)$  migre légèrement au dessus. Les pistes 2, 4 6 correspondent aux surnageants d'extraction en milieu acide, les pistes 3, 5 et 7 correspondent aux culots. La  
35 construction pBIOC706 génère une bande majoritaire reconnue par l'anticorps anti-collagène I, migrant

exactement au niveau de la bande  $\alpha 1(I)$  témoin (piste 1); la construction pBIOC707 une bande majoritaire migrant à 140 KDa. L'extraction en milieu acide semble cependant incomplète puisque les mêmes bandes se retrouvent dans  
5 les culots d'extraction correspondants (piste 3 pour pBIOC706, piste 5 pour pBIOC707). Les pistes correspondant au plant témoin (pistes 6 et 7) sont négatives.

Fig 2:

Immunotransfert, montrant quelques exemples de  
10 plants positifs transformés par les constructions PBIOC706 (plants n°40: pistes 2, 3 et n°45: pistes 4, 5) et PBIOC707 (plants n°1: pistes 6, 7; plant n°2: pistes 8, 9; plant n°14: pistes 10, 11) extraits en milieu neutre. La  
15 piste 1 présente la migration électrophorétique du collagène I témoin: la bande correspondant à  $\alpha 2(I)$  migre au niveau de la bande 116KDa des poids moléculaires standards, la bande  $\alpha 1(I)$  migre légèrement au dessus. Les  
pistes 2, 4, 6, 8, et 10 correspondent aux surnageants d'extraction en milieu neutre, les pistes 3, 5, 7, 9 et 11  
20 correspondent aux culots. Les plants 40 et 45 (construction PBIOC706) génèrent deux bandes majoritaires reconnues par l'anticorps anti-collagène I, migrant respectivement au niveau de la bande  $\alpha 1(I)$  témoin (piste  
1) et à 140 KDa. Les plants 1, 2 et 14 (construction  
25 PBIOC707) génèrent 2 bandes majoritaires migrant respectivement à 140 KDa et 160 KDa. L'extraction en milieu neutre semble complète puisque aucune bande ne se retrouve dans les culots d'extraction correspondants  
(pistes 3 et 5 pour PBIOC706, pistes 7, 9 et 11 pour  
30 PBIOC707). Les pistes correspondant au plant témoin (non montrées) sont négatives. Tous les plants positifs exhibent le même profil électrophorétique.

Fig 3:

Immunotransfert montrant les modifications des  
35 profils électrophorétiques de plants positifs transformés

par les constructions PBIOC707 (panneau A et B) et PBIOC706 (panneau C) en présence (+) ou non (-) d'agent réducteur (DTT). Les pistes 1 et 8 présentent la migration électrophorétique du collagène I témoin. Panneau A: plant  
5 n° 2 extrait en milieu acide. La bande unique après réduction (piste 2) que l'on observe après extraction en milieu acide montre une migration légèrement plus rapide en l'absence d'agent réducteur (piste 3). Ce résultat suggère que la bande observée correspond à la chaîne  $\alpha 1(I)$   
10 comprenant le N-propeptide (présence de ponts di-sulfures intrachânes). Panneau B: plant n°2 (pistes 4 et 5) et plant n°1 (pistes 6 et 7) extrait en milieu neutre. En l'absence d'agent réducteur (pistes 5 et 7) la bande supérieure, migrant à 160 KDa en présence d'agent  
15 réducteur (pistes 4 et 6), se retrouve au niveau des bandes  $\gamma$  (3 chaînes  $\alpha$  associées). La bande migrant à 140 KDa, en présence d'agent réducteur, migre légèrement plus rapidement en l'absence réducteur comme observé panneau A. Panneau C: plant n° 40 (pistes 9 et 10) et plant n°45  
20 (pistes 11 et 12) extraits en milieu neutre. Dans ce cas à nouveau, seule la migration de la bande supérieure est modifiée en l'absence de l'agent réducteur et se retrouve en majorité, comme observée panneau B, au niveau des bandes  $\gamma$ . Ces résultats indiquent que la bande supérieure  
25 pour chacune des constructions, PBIOC706 et PBIOC707, comprend le C-propeptide. Ainsi les ponts di-sulfures interchaines qui se font entre les C-propeptides de la molécule native ont pu être réalisés chez la plante. Le pontage des C-propeptides est connu pour permettre une  
30 mise en registre correcte des 3 chaînes et initier ainsi la formation de la triple hélice. Les bandes inférieures (140 KD pour la construction PBIOC707 et 120 KDa pour la construction PBIOC706) résulterait donc du clivage du domaine C-propeptidique de la chaîne  $\alpha 1(I)$  recombinante.  
35 Ce clivage est une étape importante dans la maturation du

collagène natif puisqu'il est nécessaire pour la polymérisation en fibre du collagène.

Figures 4 et 5 : représentent les fragments utilisés par les constructions pBIOC21-collagène.

5           Figure 6 :

- Partie A: Digestion par la trypsine de collagène extrait de plants transformés avec la construction pBIOC706 (plant 45) analysée par électrophorèse sur gel acrylamide 6%. Piste 1: collagène  
10 non digéré. Piste 2: collagène digéré par la trypsine à 25°C, la bande correspondant au collagène est toujours présente attestant la structure en triple hélice du collagène extrait. Piste 3: collagène dénaturé par la chaleur (50°C, 20 minutes) puis digéré par la trypsine, la  
15 bande correspondant au collagène disparaît témoin de l'efficacité de la trypsine lorsque le collagène est dénaturé.

- Partie B: Micrographie d'une réplique obtenue après ombrage tournant du collagène purifié à partir du  
20 plant 45. Les molécules de collagène sont observées sous forme de bâtonnets de 280-300 nm de long et de 1.4 nm de diamètre, caractéristique du repliement en triple hélice de tout le domaine central du collagène I. Grandissement: X 75 000.

25           Figure 7 : Construction ADEFG, pro $\alpha$ 1(I) complet : séquence protéique et nucléique

PREPARATION I : clonage d'ADNc codant pour la totalité de la chaîne pro $\alpha$ 1(I) humaine

30           1/ Préparation d'une banque ADNc MG-63

Les cellules MG-63 (N°ATCC-CRL 1427) proviennent d'un ostéosarcome humain. Environ 28 millions de cellules MG-63 cultivées en milieu DMEM (Dulbecco's modified Eagle medium, Sigma TM) sont décollées des boîtes de cultures  
35 par traitement à la trypsine 0,25% EDTA 0,05% (Ethylene

diamine tetra acetic acid). Les cellules sont centrifugées à 1000 g pendant 10 min à 4°C, et le culot cellulaire est traité avec le kit RNA-Plus™ (BIOPROBE™ Systems) qui permet de purifier les ARN totaux. Cette manipulation a  
5 été réalisée en accord avec les recommandations fournies par BIOPROBE™. Une quantité de 950 µg d'ARN totaux a ainsi été purifiée.

Les ARN poly(A)+ ont été séparés par chromatographie sur colonne oligo-dT cellulose (Boehringer  
10 Mannheim) en suivant la technique décrite par (Aviv et al 1972). Une suspension de soude 0,1 N contenant 0,2 g d'oligo dT cellulose est déposée dans une colonne. Cette colonne est alors lavée par 3 ml d'eau milli-Q stérile, puis par 15 ml d'un tampon porteur 1X (Tris-Cl 20 mM, pH  
15 7,6; NaCl 0,5 M; EDTA 1 mM, pH 8; sodium lauryl sarcosinate 0,1%). L'effluent de la colonne ainsi équilibrée est alors à un pH inférieure à 8.

A une solution d'ARN totaux (950 µg) incubée 5 min à 65°C est ajouté un volume de tampon porteur 2X  
20 (Tris-Cl 40 mM, pH 7,6; NaCl 1 M; EDTA 2 mM, pH 8; sodium lauryl sarcosinate 0,2%). Ce mélange est refroidi à température ambiante avant d'être déposé dans la colonne oligo dT cellulose. L'éluat est récupéré, incubé 5 min à 65 °C, refroidi à température ambiante et repassé dans la  
25 colonne oligo-dT cellulose. La colonne est alors lavée par 10 ml de tampon porteur 1X (Tris-Cl 20 mM, pH 7,6; NaCl 0,5 M; EDTA 1 mM, pH 8; sodium lauryl sarcosinate 0,1%), ce qui permet l'élution des ARN poly(A)-, les ARN poly(A)+ restant fixés à la matrice dT cellulose par leur extrémité  
30 3' poly-adenylée. A la colonne oligo dT-ARN poly(A)+ sont alors ajoutés 2 ml de tampon de décrochage (Tris-Cl 10 mM, pH 7,6; EDTA 1 mM, pH 8; Lauryl sulfate 0,05%), et l'éluat d'ARN poly(A)+ est ajusté à 0,3 M en acétate de sodium pH 5,2. A cette solution sont apportés 2,2 volumes d'éthanol  
35 100, et les ARN poly(A)+ sont précipités pendant 24 heures à - 20°C.

Après précipitation, la solution est centrifugée à 12000 g pendant 30 min à 4°C. Le culot d'ARN poly(A)+ est lavé par 10 ml d'éthanol 70, lyophilisé et repris par 50 µl d'eau milli-Q stérile. La concentration de cette  
5 fraction d'ARN poly(A)+ est de 1,2 µg/µl.

Pour réaliser la banque ADNc, nous avons utilisé des kits commercialisés par Amersham. Ce sont les kits:

- cDNA synthesis system plus, RPN 1256
- cDNA rapid adaptator ligation module, RPN 1712
- 10 - cDNA rapid cloning module, RPN 1716
- lambda-DNA *in vitro* packaging module, RPN 1717

Chaque étape est réalisée en suivant le protocole fourni par la fabricant.

La solution d'ARN poly(A)+ est incubée 5 min à  
15 65°C, puis refroidie dans la glace afin d'éliminer les structures secondaires présentes dans certains ARN. Cette solution est divisée en deux fractions, et après appariement soit d'amorce dT (12-18 mer), soit d'amorces aléatoires (hexanucléotides) avec les ARN poly(A)+, la  
20 synthèse du premier brin ADN (complémentaire à la matrice ARN) va être réalisée par la reverse transcriptase. Au mélange hybride ARN/ADN est ajouté la ribonucléase H et l'ADN polymérase I d'*Escherichia coli*. La ribonucléase H permet de couper aléatoirement le brin ARN dans un  
25 complexe ARN/ADN. L'ADN polymérase ajoute des nucléotides aux extrémités 3'-OH libres créées par la ribonucléase H, grâce à son activité polymérase 5'-3'. L'activité exonucléase 5'-3' de l'ADN polymérase I permet d'éliminer le ribonucléotide en 5' et de poursuivre la synthèse du  
30 second brin ADNc. Après inactivation des enzymes par traitement à 70°C pendant 10 min, l'addition de T4 DNA polymérase permet d'obtenir des ADN double brin à extrémités franches grâce à son activité exonucléase 3'-5'. Des adaptateurs EcoRI sont alors ligués aux extrémités  
35 franches des ADNc (ADN complémentaires) . Les adaptateurs libres sont éliminés sur une colonne fournie par Amersham TM , et les ADNc adaptés en EcoRI sont ligués dans le site



EcoRI du vecteur lambda MOSElox. Après ligation, une encapsidation in vitro est réalisée, ce qui est la dernière étape avant l'obtention de la banque ADNc. Environ 800000 clones ADNc indépendants ont ainsi été  
5 obtenus.

2/ clonage d' ADNc codant pour la chaîne pro $\alpha$ 1(I) humaine.

a/Préparation de bactéries ER1647 aptes à être infectées par les phages

10 Cette souche bactérienne est recD-, mrc A- (Amersham TM). Ces bactéries sont maintenues dans des boîtes de L-broth/agar 1,5% (L-broth: pour 1 litre, 10 g de tryptone, 5 g d'extraits de levure, 10 g de NaCl) supplémentée en tétracycline (50  $\mu$ g/ml). Une colonie  
15 ER1647 est ensemencée dans 10 ml de L-broth/tétracycline (50  $\mu$ g/ml), et la culture est réalisée sous agitation à 37°C pendant 8 heures. La culture ER1647 est ensuite centrifugée à 4°C pendant 10 min à 3000 g, et le culot bactérien est resuspendu dans 2 ml de MgSO<sub>4</sub> 10 mM froid.

20 b/ préparation de répliques de nitrocellulose

Pour le criblage, 4 boîtes (12X12 cm) L-broth/agar 1,5% de 25000 clones ADNc indépendants ont été réalisées. Pour cela, à 250  $\mu$ l de bactéries ER1647 traitées au MgSO<sub>4</sub> 10 mM, ont été apportés 25000 phages de  
25 la banque ADNc. L'ensemble est incubé à 37°C pendant 15 min, temps nécessaire à une infection efficace des bactéries. Après apport de 7 ml de L-broth/agarose medium 0,7% (Sigma TM) maintenu à l'état liquide à 45°C, l'ensemble est déposé sur les boîtes de culture (12X12  
30 cm).

Ces boîtes sont incubées à 37°C pendant 7-8 heures, temps nécessaire à l'apparition des plages de lyse. Elles sont alors utilisables pour le transfert de l'ADN des plages de lyse sur des répliques de  
35 nitrocellulose. La nitrocellulose utilisée est Hybond-C d'Amersham TM. Pour chaque boîte, deux répliques sont

réalisées par simple contact plages de lyses/réplique de nitrocellulose. Le temps de contact est de 2 min pour la première réplique, et de 4 min pour la seconde. L'ADN de phages présents sur les répliques est dénaturé pendant 4 min dans une solution de soude 0,5 M, NaCl 1,5 M. La soude est alors neutralisée par deux bains successifs de 3 min dans une solution NaCl 1,5 M, Tris-Cl 0,5 M pH 7. Les répliques de nitrocellulose sont rincées dans une solution de 2X SSC (NaCl 0,3 M, citrate trisodique 30 mM), séchées sur papier Whatman 3MM, et l'ADN dénaturé de phages est fixé sur les répliques de nitrocellulose par à une incubation de 2 heures à 80°C.

c/ préparation de sondes spécifiques du collagène de type I humain

Deux oligonucléotides de synthèse ont été utilisés pour obtenir des sondes spécifiques de la chaîne pro $\alpha$ 1(I).

Oligonucléotides BIOC5, 5'-CTCGGGTTTCCACACGT-3', séquence complémentaire aux nucléotides 152 à 178 spécifique de la chaîne pro $\alpha$ 1(I).

Oligonucléotide BIOC7, 5'-GCAAGACAGTGATTGAA-3', séquence 4268 à 4274 spécifique de la chaîne pro $\alpha$ 1(I).

Ces oligonucléotides ont été marqué au 32P, en utilisant du gamma P32ATP à 3000 Ci/mmol (Amersham TM) et de la T4 polynucléotide Kinase (Promega).

Conditions de marquage,	50 ng BIOC5 ou BIOC7	10 $\mu$ l
	H2O milli-Q	4,5 $\mu$ l
	gamma32PATP, 50 $\mu$ CI	5 $\mu$ l
	T4 Polynucléotide kinase	
	(10 u/ $\mu$ l; Promega TM)	0,5 $\mu$ l

Ce mélange est incubé pendant 30 min à 37°C. La réaction est stoppée par ajout de 1  $\mu$ l EDTA 0,5 M, pH 8.

d/ hybridation des répliques de nitrocellulose

Les répliques de nitrocellulose sont déposées dans 50 ml d'une solution de préhybridation, et incubées à 50°C pendant 15 min.

Solution de préhybridation:

- 5                   6XSSC (NaCl 0,3 M, citrate trisodique 90 mM)  
                  Lauryl sulfate 0,1%  
                  Sodium pyrophosphate 0,05%  
                  polyvinylpyrrolidone 0,02%  
                  albumine sérique bovine 0,02%  
10                  Ficoll 400 0,02%  
                  ADN dénaturé de sperme de saumon,  
                  à 100 µg/ml (Boehringer Mannheim)

- Les deux sondes BIO5 et BIO7 sont alors ajoutées, et l'hybridation est réalisée pendant 2 heures à  
15   45°C. Après hybridation, les répliques de nitrocellulose subissent trois bains successifs de 15 min afin d'éliminer les hybridations aspécifiques.

- Premier bain: 6XSSC (NaCl 0,9 M, citrate trisodique 90 mM), sodium pyrophosphate 0,05%, 15 min à  
20   température ambiante.

                  Second bain: 3XSSC (NaCl 0,45 M, citrate trisodique 45 mM), sodium pyrophosphate 0,05%, 15 min à température ambiante.

- Troisième bain: 6XSSC (NaCl 0,9 M, citrate trisodique 90 mM), sodium pyrophosphate 0,05%, 15 min à  
25   45°C.

- Les filtres sont alors déposés sur un plastique, entouré d'un film de plastique transparent et soumis à autoradiographie à -80°C pendant 24 heures en présence  
30   d'un film.

- Cette manipulation nous a permis d'obtenir plusieurs clones positifs en duplica. Chaque plage de lyse positive dans les deux répliques est récupérée à l'aide d'une pipette pasteur, et déposée dans 1 ml de tampon SM  
35   (NaCl 100mM, MgSO4 8 mM, Tris-Cl 50 mM pH 7,5, gélatine 0,01%).

Un second criblage nous a permis de purifier les différents clones positifs, les conditions d'hybridation et de lavages des répliques de nitrocellulose étant les mêmes que pour le premier criblage. Les clones purifiés  
5 lors de ce second criblage sont prélevés et déposés dans 500 µl de SM.

e/ Sous clonage automatique des clones ADNc en plasmide par le système cre-lox.

Le vecteur bactériophagique lambda MOSElox  
10 comporte un plasmide interne flanqué à ses deux extrémités par les sites lox de 34 pb, dérivés du bactériophage P1. Ainsi, des sous-clones plasmidiens peuvent être obtenus en infectant une souche bactérienne exprimant la cre recombinase P1, qui reconnaît les séquences lox entourant  
15 le plasmide interne au bactériophage lambda MOSElox et qui permettra l'excision du plasmide par recombinaison spécifique au niveau des séquences lox.

Dans notre cas, nous avons utilisé la souche bactérienne BM25.8 qui est lysogène pour P1 et lambda-imm434kan, ces deux plasmides étant maintenus en ajoutant  
20 au milieu de croissance de la souche BMB25.8, de la kanamycine (50 µg/ml) et du chloramphénicol (50 µg/ml).

En pratique, 10 ml de L-broth complétés en kanamycine et chloramphénicol (50 µg/ml) sontensemencés  
25 avec les bactéries BM25.8 et incubés à 37°C sous agitation. Lorsque la D0600 de cette culture est de 0,9, l'incubation est stoppée et les bactéries peuvent être conservées pendant 3 jours à 4°C.

Une fraction aliquote de la suspension phagique  
30 des clones ADNc positifs (10 µl) sont mis en contact avec 200 µl de bactéries BM25.8, et le mélange incubé à 37°C pendant 15 min afin d'obtenir une infection efficace. Le mélange est alors déposé dans une boîte L-Broth/agar 1,5% supplémentée en ampicilline (50 µg/ml), et les boîtes sont  
35 incubées à 37°C pendant 12 heures. La présence d'ampicilline permet de sélectionner les bactéries ayant intégré un bactériophage, et plus spécialement celles où

s'est produite la recombinaison spécifique cre-lox, le plasmide résultant possédant un gène de résistance à l'ampicilline, et l'insert ADNc au niveau du site de clonage EcoRI.

5 f/ caractérisation des clones ADNc alpha22 et 1alpha3

Deux des clones ADN complémentaires positifs ont été utilisés ultérieurement. Ce sont les clones ADNc alpha 22 et 1alpha3.

10 En vue de séquencer ces clones, 3 ml de L-broth supplémenté en ampicilline (50 µg/ml) ont été ensemencés soit par alpha22, soit par 1alpha3. Ces mélanges ont été incubés sous agitation à 37°C pendant une nuit.

Pour purifier l'ADN plasmidien, la culture est  
15 tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), qui permet de purifier de l'ADN plasmidien super enroulé. Cette manipulation est réalisée suivant le protocole fourni par le fabricant. Environ 10 µg d'ADN  
20 plasmidien est purifié grâce à ce kit, lorsque le plasmide utilisé est le plasmide MOSElox.

La dernière étape est de séquencer les deux extrémités des clones ADNc alpha22 et 1alpha3, afin de déterminer quelles informations ils comportent.

25 Pour réaliser 2 réactions de séquence, 40 µl d'ADN, soit 7-8 µg, purifié par le système Wizard sont dénaturés par ajout de 10 µl de NaOH 2N et incubés 10 min à température ambiante. Après neutralisation par apport de 12 µl d'acétate de sodium 3 M pH 5,2, l'ADN monocaténaire  
30 est précipité par apport de 2,5 volume d'éthanol 100 et incubation à -20°C pendant 30 min. Après précipitation, l'ADN est centrifugé à 4°C pendant 10 min à 12 000 g, lavé par 1 ml d'éthanol 70, lyophilisé et repris par 20 µl d'eau milli-Q.

35 Les réactions de séquence ont été réalisées avec le kit "T7 sequencing" (Pharmacia) qui utilise la technique de terminaison de chaînes et la T7 DNA

polymerase. Le radio-élément est le [ $\alpha$ -35S]-dATP (Amersham TM) à 1350 Ci/mmol. Les réactions de séquences ont été réalisées en suivant les recommandation du fabricant.

Principe. Après appariement d'une amorce sur une  
5 matrice d'ADN monocaténaire, une réaction d'élongation est réalisée par la T7 DNA polymérase en présence de dGTP, dTTP, dCTP et de 35-SdATP. L'apport d'un excès des 4 dNTP froids, ainsi que du ddATP (dideoxyadenosine triphosphate) ou du ddGTP ou du ddCTP ou du ddTTP, permet de réaliser  
10 séparément 4 réactions de terminaison. Lors de l'incorporation d'un ddNTP, l'absence du groupement hydroxyl bloque la réaction d'élongation.

Dans notre cas, les amorces utilisées sont les oligonucléotides T7 gene 10 et T7 terminator (Amersham TM)  
15 qui se situent de part et d'autre des clones ADNc alpha22 et lalpha3..

T7 gene primer, 5'- TGAGGTTGTAGAAGTTCCG-3'

T7 terminator primer, 5'- GCTAGTTATTGCTCAGCGG-3'

L'analyse des séquences est réalisée sur gel  
20 polyacrylamide-urée (6%-7M).

Gel de séquence

La solution de départ est: 39,6 g d'urée, 8,25 ml de TBE 10X (Tris borate 0,9 M, EDTA 20 mM, pH 8), 12,32 ml d'acrylamide 40% (38 g d'acrylamide, 2 g de bis-  
25 acrylamide, qsp 82,5 ml eau milli-Q).

Après dissolution de l'urée, 400  $\mu$ l de persulfate d'ammonium 10% et 66  $\mu$ l de TEMED (N, N, N', N'-tétraméthyl-éthylène diamine) sont ajoutés et le gel est coulé. Après polymérisation, le gel est préchauffé pendant  
30 30 min à 60 kW. Les échantillons sont alors déposés et soumis à séparation sous une puissance de 60 kW. Après électrophorèse, le gel est transféré sur du papier Whatman 3MM, séché pendant 30 min sous vide à 80°C, et soumis à autoradiographie directe à température ambiante avec un .  
35 La lecture des séquences nous a permis de confirmer que

les clones ADNC alpha22 et 1alpha3 codent pour la chaîne pro $\alpha$ 1(I) humaine.

Plus précisément, le clone ADNC 1alpha 3 comporte 83 pb de la partie 5' non traduite et les 1920 premières bases codant pour la chaîne pro $\alpha$ 1(I) humaine. Le clone ADNC Alpha 22 comporte la séquence codant pour les acides aminés 171 à 1454 de la chaîne pro $\alpha$ 1(I) humaine et environ 500 pb de la région 3' non-traduite.

La conformité du gène obtenu est vérifiée par rapport aux séquences décrites (LI S-W, 1994)

PREPARATION II : clonage de fragments d'ADN nécessaires à la réalisation des constructions selon l'invention.

1/ Fragment A

Le fragment A possède 4 nucléotides (-4 à -1) en amont du codon initiateur de traduction ATG, la totalité de la séquence codant pour le signal peptide (nucléotides 1 à 66) et pour le domaine amino-propeptide (nucléotides 67 à 479) de la chaîne pro $\alpha$ 1(I) humaine. Les nucléotides 474, 475 et 477, CTT, sont substitués par les bases GCA, ce qui permet la création d'un site de restriction NheI (GCTAGC) de la position 474 à 479. Cette création du site NheI entraîne la substitution des acides aminés Asn et Phe par une lysine et une leucine en position 158 et 159 de la chaîne pro $\alpha$ 1(I) humaine.

Amplification du fragment A

Deux oligonucléotides de synthèse ont été utilisés.

Oligonucléotide sens, BIOC85,  
5' TATCCGCGGAAGCTTAGACATGTTCAGCTTTGTGGACCTCCGGCTC  
CTGC-3'

Cet oligonucléotide comporte en 5' (nucléotides soulignés) les séquences permettant la création de deux

sites de restriction, SacII (CCGCGG) et HindIII (AAGCTT) qui seront utilisés comme sites de clonage. A la suite de ces deux sites sont inclus les nucléotides -4 à -1 et 1 à 31 spécifiques de la séquence codante de la chaîne pro $\alpha$ (I) humaine.

Oligonucléotide antisens, BIOC83,

5' TATTCTAGAGCTAGCTTTCCTCCGAGGCCAGGGGGTCCGGGAGGT-3'

Cet oligonucléotide présente en 5' (nucléotides soulignés) les séquences permettant la création des sites de restriction XbaI (TCTAGA) et NheI (GCTAGC). Le site XbaI sera utilisé pour le clonage du fragment A, alors que le site NheI, décrit dans la partie information, sert à délimiter la séquence codant pour le domaine aminopropeptide. Suite à ces deux sites est retrouvée la séquence complémentaire aux nucléotides 445 à 474 de la séquence codante de la chaîne pro $\alpha$ 1(I) humaine.

La matrice ADN utilisée est le clone ADNC 1alpha3 qui comporte 83 pb de la partie 5' non traduite et les 1920 premières bases codant pour la chaîne pro $\alpha$ 1(I) humaine.

Pour obtenir le fragment A, les conditions d'amplification par PCR (Polymerase Chain Reaction) ont été:

25 cycles:	Dénaturation,	30s à 94°C
	Appariement,	30s à 55°C
	Elongation,	30s à 72°C.

A la fin des 25 cycles, une extension supplémentaire de 5 min à 72°C est réalisée.

Solution de départ:

75 mM Tris-Cl pH 9,0 (à 25°C)
20 mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>
0,01% (Poids : volume) Tween 20
1,5 mM MgCl <sub>2</sub>
BIOC85, 0,5 $\mu$ M
BIOC83, 0,5 $\mu$ M



chaque dNTP, 0,2 mM  
matrice, 1 ng du clone ADNc 1alpha3  
0,25 unité de DNA polymerase)

Clonage du fragment A

5           Après amplification, la solution est extraite  
par un volume de phénol/chloroforme/alcool isoamylique  
(rapport 50/48/2). Après deux minutes de centrifugation à  
10 000 g, la phase aqueuse est reprise et extraite par un  
volume de chloroforme/alcool isoamylique (rapport 24/1).  
10   Après une centrifugation de quelques secondes à 10 000 g,  
la phase aqueuse est reprise, ajustée à une concentration  
en NaCl de 0,2 M et deux volumes d'alcool 100 sont  
ajoutés. L'ADN est précipité pendant 2 heures à -20°C,  
puis est centrifugé à 4°C pendant 10 min à 12 000 g. Le  
15   culot d'ADN est lavé avec 1 ml d'éthanol 70, lyophilisé,  
et resuspendu dans 34 µl d'eau milli-Q stérile.

          A cette solution d'ADN sont ajoutés 4 µl de  
tampon de digestion C (Promega TM), 1 µl de l'endonucléase  
de restriction SacII (10 u/µl; Promega TM) et 1 µl  
20   d'enzyme XbaI (12 u/µl; Promega). La digestion est  
réalisée pendant 2 heures à 37°C.

          Suite à la digestion enzymatique, 8 µl de tampon  
stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés  
à la solution d'ADN qui est analysée par électrophorèse  
25   sur gel d'agarose 2% à bas point de fusion Nu Sieve-GTG  
(FMC) à un voltage constant de 60 volts. Le tampon  
d'électrophorèse est du TBE 0,5X (Tris borate 45 mM, EDTA  
(Acide éthylène-diamine-tétra-acétique) 1mM, pH 8). Après  
migration, le gel est coloré pendant 15 min dans une  
30   solution de TBE 0,5X contenant 40 ng/ml de bromure  
d'ethidium.

          La bande correspondant au fragment A est  
découpée sous UV, déposée dans 350 µl d'une solution Tris  
20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à  
35   65°C. Après dissolution de l'agarose, l'échantillon est  
extraît par un volume de phénol et centrifugé pendant 5  
min à 10 000 g. La phase aqueuse est reprise, extraite par

un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) et centrifugée pendant 2 min à 10 000 g. La phase aqueuse est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques  
5 secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en NaCl de 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

Après centrifugation à 4°C pendant 10 min à 12  
10 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a  
15 permis d'estimer la concentration de la solution purifiée du fragment A SacII/XbaI à 30 ng/µl.

Le plasmide utilisé pour cloner le fragment A est le vecteur pBluescript II SK(+) commercialisé par Stratagene TM. Ce plasmide comporte un gène de résistance  
20 à l'ampicilline.

Une quantité de 200 ng du plasmide pBluescript II SK(+) est soumise à une digestion par les endonucléases de restriction SacII et XbaI. Cette digestion est réalisée dans un volume de 20 µl en présence de 2 µl de tampon C  
25 (Promega TM), 1 µl d'enzyme SacII (10 u/µl; Promega TM) et 1 µl d'enzyme XbaI (12 u/µl; Promega TM), pendant 2 heures à 37°C. Après digestion, 2 µl de tampon porteur (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés au mélange et analysé sur gel d'agarose Seaplaque GTG (FMC,  
30 Rockland) 0,8% à un voltage constant de 60 volts, le tampon d'électrophorèse étant du TBE 0,5X. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'ethidium à 40 ng/ml. La bande d'ADN correspondant au pBluescript II SK(+) linéarisé par SacII  
35 et XbaI est prélevée et déposée dans 350 µl d'une solution de Tris 20mM pH 8, EDTA 1 mM pH 8. Cet échantillon est incubé pendant 5 min à 65°C, extrait par un volume de

phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) et centrifugée pendant 2 min à 10 000 g. La phase aqueuse  
5 est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en NaCl de 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

10 Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 10 ng/µl du plasmide pBluescript II SK(+) SacII/XbaI.

15 La ligation du fragment A (SacII/XbaI) au plasmide pBluescript II SK(+) (SacII/XbaI) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5 alpha (Gibco BRL,  
20 Paris). Le génotype de cette souche est: supE 44 Δ-lacU 169(Ø80 lacZΔM15) hsdR 17 recA 1 endA 1 gyrA 96 thi- 1 relA 1. Pour obtenir des bactéries compétentes, 5 ml d'une culture en phase exponentielle (DO600 de 0.7) sont centrifugés à 1500 g pendant 15 min à 4°C. Le culot  
25 bactérien est alors resuspendu dans 500 µl de CaCl<sub>2</sub> 100 mM froid. Pour la transformation, 10 µl du mélange de ligation est mélangé à 200 µl de bactéries DH5 alpha compétentes et l'échantillon est placé à 4°C pendant 30 min. Après un choc thermique à 42°C pendant 40 s, 500 µl  
30 de L-broth (10 g de tryptone, 5 g d'extrait de levure, 10 g de NaCl, par litre de milieu) sont ajoutés, et le mélange est incubé pendant 30 min à 37°C.

Ce mélange ainsi que 20 µl d'IPTG (Isopropylthio-beta-D-galactoside; 200mM) et 20 µl de X-Gal (4 chloro-3 indolyl-beta-D galactoside; 20 mg/ml) sont  
35 ensuite déposés sur une boîte de L-Broth-agar (L-Broth

plus agar 1.5%) supplémentée en ampicilline (50 µg/ml), et la boîte est incubée pendant 15 heures à 37°C. L'apport d'IPTG et de X-Gal permet l'utilisation du test de coloration qui permet de reconnaître les plasmides ayant  
5 incorporé un insert.

En effet, la souche bactérienne DH5 alpha est Lac-, et produit une forme non fonctionnelle de la beta-galactosidase. Le vecteur pBluescript II SK(+) possède un segment d'ADN de l'opéron lactose d'*Escherichia coli* qui  
10 code pour la partie amino-terminale de cette enzyme. L'induction de synthèse par l'IPTG permet une complémentation alpha qui restaure l'activité de cette galactosidase. En présence d'un analogue du galactose, le X-Gal, les bactéries produisent des pigments bleus qui  
15 entraînent la formation d'une colonie bactérienne colorée. Le vecteur pBluescript II SK(+) possède une région de multiples clonages (comprenant les sites SacII/XbaI) située dans le gène LacZ. L'insertion du fragment A au sein de cette région de clonage, abolit la complémentation  
20 alpha, ce qui entraîne l'apparition de colonies bactériennes blanches.

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces  
25 cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard  
30 (Promega), qui permet de purifier de l'ADN plasmidien super enroulé. Cette manipulation est réalisée suivant le protocole fourni par le fabricant. Environ 10 µg d'ADN plasmidien est purifié grâce à ce kit, lorsque le plasmide utilisé est le pBluescript II SK(+).

35 Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par les endonucléases de restriction SacII et XbaI, afin de

vérifier le clonage du fragment A. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment A est vérifiée sous UV.

- 5 La dernière étape est de vérifier l'intégrité de la séquence du fragment A, la technique d'amplification par PCR (polymerase chain reaction) pouvant entraîner des mutations.

10 Les réactions de séquence réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase (cf PREPARATION 1) permet de valider l'intégrité de la séquence du fragment A.

#### 2/ Fragment B

- 15 Le fragment B possède 3 bases (TAA) en amont de la séquence codant pour le signal peptide PRS (PATHOGENESIS-RELATED PROTEIN S) (nucléotides 1 à 75), et les bases 67 à 77 de la séquence codant pour la partie amino-terminale du domaine amino-propeptide de la chaîne
- 20 pro $\alpha$ 1(I) humaine. Les nucléotides 73, 74 et 77, GAG, sont substitués par les bases CTC, ce qui permet la création d'un site de restriction NheI (GCTAGC) de la position 72 à 77. Cette création du site NheI entraîne la substitution des acides aminés Glu et Gly par une leucine et une
- 25 alanine en position 25 et 26 de la chaîne pro $\alpha$ 1(I) humaine.

#### Amplification du fragment B

Trois oligonucléotides de synthèse ( BIOC95 et BIOC93 et 045) ont été utilisés.

- 30 Oligonucléotide matrice, 045,  
5'-TAAATGAAGTTCCTCAAAAGTTTCCCTTTTATGCCTTCCTT  
TGTTTTGGCCAATACTTTGTAGCTGTTACTCATGCTGCC-3'

Cet oligonucléotide comporte en gras la séquence correspondant à celle du peptide signal de PRS  
35 (PATHOGENESIS-RELATED PROTEIN S). Il servira de matrice

ADN lors de l'amplification par PCR (Polymerase Chain Reaction) du fragment B

Oligonucléotide sens, BIOC95,

5'-TATCCGCGGAAGCTTTAAATGAACTTCCTCAAAAGTTCCCC-3'

5 Cet oligonucléotide comporte en 5' (nucléotides soulignés) les séquences permettant la création de deux sites de restriction, SacII (CCGCGG) et HindIII (AAGCTT) qui seront utilisés comme sites de clonage. A la suite de ces deux sites sont inclus les nucléotides -3 à -1 et 1 à 10 24 spécifiques de la séquence codant pour le signal peptide de PRS (PATHOGENIC-RELATED PROTEIN S).

Oligonucléotide antisens, BIOC93,

5'- ATGCTAGCTCTTGAGCATGAGTAACAGCTACAAAGTA-3'

15 Cet oligonucléotide présente en 5' (nucléotides soulignés) la séquence permettant la création du site de restriction NheI (GCTAGC). Le site NheI, décrit dans la partie information, sert à délimiter la séquence codant pour le domaine amino-propeptide. Après le site NheI est retrouvée la séquence complémentaire aux nucléotides 67 à 20 71 de la séquence codante de la chaîne pro $\alpha$ 1(I) humaine et 52 à 75 de la séquence codant pour le signal peptide de PRS (PATHOGENESIS-RELATED PROTEIN S).

#### Clonage du fragment B

Après l'amplification PCR, la solution est 25 extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2), comme décrit dans préparation II1. Le culot d'ADN est lavé avec 1 ml d'éthanol 70, lyophilisé, et resuspendu dans 40  $\mu$ l d'eau milli-Q stérile.

30 Une fraction aliquote de 4  $\mu$ l, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment B à 20 ng/ $\mu$ l.

35 Le plasmide utilisé pour cloner le fragment B est le vecteur pBluescript II SK(+) commercialisé par

Stratagene. Ce plasmide comporte un gène de résistance à l'ampicilline.

Une quantité de 200 ng du plasmide pBluescript II SK(+) est soumise à une digestion par l'endonucléase de restriction EcoRV (GATATC), qui permet d'obtenir un vecteur linéarisé à extrémités franches. Après digestion, 80 µl de TE (Tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés et la solution est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans préparation III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 10 ng/µl du plasmide pBluescript II SK(+) EcoRV.

La ligation du fragment B au plasmide pBluescript II SK(+) (EcoRV) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5 alpha (Gibco BRL, Paris) cf ci-dessus.

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par les endonucléases de restriction BamHI et HindIII, afin de vérifier le clonage du fragment B. Les sites de restriction BamHI et HindIII sont localisés de part et d'autre du site de clonage EcoRV. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure

d'éthidium à 40 ng/ml, et la présence du fragment B est vérifiée sous UV.

Les réactions de séquence réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymérase (cf PREPARATION 1) permet de valider l'intégrité de la séquence du fragment B.

### 3/ Fragment C

Le fragment C possède la quasi totalité de la séquence codant pour le domaine amino-propeptide (nucléotides 72 à 479) de la chaîne pro $\alpha$ 1(I) humaine. Les nucléotides 73, 74 et 77, GAG, sont substitués par les bases CTC, ce qui permet la création d'un site de restriction NheI (GCTAGC) de la position 72 à 77. Cette création du site NheI entraîne la substitution des acides aminés Glu et Gly par une leucine et une alanine en position 25 et 26 de la chaîne pro $\alpha$ 1(I) humaine. Les nucléotides 474, 475 et 477, CTT, sont substitués par les bases GCA, ce qui permet la création d'un site de restriction NheI (GCTAGC) de la position 474 à 479. Cette création du site NheI entraîne la substitution des acides aminés Asn et Phe par une lysine et une leucine en position 158 et 159 de la chaîne pro $\alpha$ 1(I) humaine.

### Amplification du fragment C

Deux oligonucléotides de synthèse ont été utilisés.

Oligonucléotide sens, BIOC855,

5'- ATGCTAGCCCAAGTCGAGGGCCAAGACGAAG-3'

Cet oligonucléotide comporte en 5' (nucléotides soulignés) la séquence permettant la création d'un site de restriction NheI (GCTAGC) qui sera utilisé comme site de clonage. Le site NheI, décrit dans la partie information, sert à délimiter la séquence codant pour le domaine amino-propeptide (extrémité amino-terminale). A la suite de ce



site sont inclus les nucléotides 78 à 100 spécifiques de la séquence codante de la chaîne pro $\alpha$ 1(I) humaine.

Oligonucléotide antisens, BIOC83,

5' TAT TCTAGAGCTAGCT TTCCTCCGAGGCCAGGGGGTCCGGGA

5 GGT-3'

Cet oligonucléotide présente en 5' (nucléotides soulignés) les séquences permettant la création des sites de restriction XbaI (TCTAGA) et NheI (GCTAGC). Le site XbaI utilisé lors du clonage du fragment A, ne sera pas  
10 utilisé dans le cadre du fragment C. Le site NheI, décrit dans la partie information, sert à délimiter la séquence codant pour le domaine amino-propeptide (extrémité carboxy-terminale). Suite à ces deux sites est retrouvée la séquence complémentaire aux nucléotides 445 à 474 de la  
15 séquence codante de la chaîne pro $\alpha$ 1(I) humaine.

La matrice ADN utilisée est le clone ADNc 1alpha3 qui comporte 83 pb de la partie 5' non traduite et les 1920 premières bases codant pour la chaîne pro $\alpha$ 1(I) humaine.

20 Clonage du fragment C

Après amplification, la solution est extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) comme décrit dans préparation III. Le culot d'ADN est lavé avec 1 ml d'éthanol 70, lyophilisé,  
25 et resuspendu dans 40  $\mu$ l d'eau milli-Q stérile.

Une fraction aliquote de 4  $\mu$ l, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a permis d'estimer la concentration de la  
30 solution purifiée du fragment C à 20 ng/ $\mu$ l.

Le plasmide utilisé pour cloner le fragment C est le vecteur pBluescript II SK(+) commercialisé par Stratagene TM (La jolla). Ce plasmide comporte un gène de résistance à l'ampicilline.

Une quantité de 200 ng du plasmide pBluescript II SK(+) est soumise à une digestion par l'endonucléase de restriction EcoRV (GATATC), qui permet d'obtenir un vecteur linéarisé à extrémités franches. Après digestion, 5 80 µl de TE (Tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés et la solution est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 10 ng/µl du plasmide pBluescript II SK(+) EcoRV.

La ligation du fragment C au plasmide pBluescript II SK(+) (EcoRV) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5 alpha (Gibco BRL, Paris).

20 Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

25 Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus) 30 est soumise à une digestion enzymatique par les endonucléases de restriction BamHI et HindIII, afin de vérifier le clonage du fragment C. Les sites de restriction BamHI et HindIII sont localisés de part et d'autre du site de clonage EcoRV. Après électrophorèse, le 35 gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment C est vérifiée sous UV.

Les réactions de séquence réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase (cf PREPARATION 1) permet de valider l'intégrité de la séquence du fragment C.

#### 4/ Fragment D

Le fragment D comporte la totalité de la séquence codant pour le domaine amino-télopeptide (nucléotides 474 à 534) et la séquence codant pour la partie amino-terminale de la région hélicoïdale (nucléotides 535 à 1920) de la chaîne pro $\alpha$ 1(I) humaine.

Les nucléotides 474, 475 et 477, CTT, sont substitués par les bases GCA, ce qui permet la création d'un site de restriction NheI (GCTAGC) de la position 474 à 479. Cette création du site NheI entraîne la substitution des acides aminés Asn et Phe par une lysine et une leucine en position 158 et 159 de la chaîne pro $\alpha$ 1(I) humaine. Un site de restriction DraIII est localisé de la position 1709 à 1717 (CACCTGGTG), alors qu'un site BamHI dû au plasmide lambda MoSElox (Amersham TM) est situé en extrémité 3'.

#### Amplification du fragment D

Deux oligonucléotides de synthèse ont été utilisés.

Oligonucléotide sens, BIOC65,

5'- TATTCTAGAGCTAGCTCCCCAGCTGTCTTATGGCTATGATGAG-  
3'

Cet oligonucléotide présente en 5' (nucléotides soulignés) les séquences permettant la création des sites de restriction XbaI (TCTAGA) et NheI (GCTAGC). Le site XbaI sera utilisé pour le clonage du fragment D, alors que le site NheI, décrit dans la partie information, sert à délimiter la séquence codant pour le domaine amino-propeptide. Après ces deux sites est retrouvée la séquence (nucléotides 480 à 507) codant pour le début du domaine amino-télopeptide de la chaîne pro $\alpha$ 1(I) humaine.

Oligonucléotide antisens, T7 gene 10 primer,  
5'- CTGAGGTTGTAGAAGTTCCG-3'

Cet oligonucléotide est localisé en juste en  
aval de l'extrémité 3' du clone ADNc 1alpha3, et d'un site  
5 de restriction BamHI (GGATCC) qui sera utilisé pour le  
clonage du fragment D.

La matrice ADN utilisée est le clone ADNc  
1alpha3 qui comporte 83 pb de la partie 5' non traduite et  
les 1920 premières bases codant pour la chaîne pro $\alpha$ 1(I)  
10 humaine.

#### Clonage du fragment D

Après amplification, la solution est extraite  
par un volume de phénol/chloroforme/alcool isoamylique  
(rapport 50/48/2) comme dans PREPARATION II1. Le culot  
15 d'ADN est lavé avec 1 ml d'éthanol 70, lyophilisé, et  
resuspendu dans 34  $\mu$ l d'eau milli-Q stérile.

La digestion par Bam H1 Xba1 est réalisée  
pendant 2 heures à 37°C dans un volume final de 40  $\mu$ l.

Suite à la digestion enzymatique, 8  $\mu$ l de tampon  
20 stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés  
à la solution d'ADN qui est analysée par électrophorèse  
sur gel d'agarose 0,8% à bas point de fusion SeaPlaque-GTG  
(FMC, Rockland) à un voltage constant de 60 volts. Le  
tampon d'électrophorèse est du TBE 0,5X (Tris borate 45 mM,  
25 EDTA (Acide éthylène-diamine-tétra-acétique) 1mM, pH 8).  
Après migration, le gel est coloré pendant 15 min dans une  
solution de TBE 0,5X contenant 40 ng/ml de bromure  
d'ethidium.

La bande correspondant au fragment D est  
30 découpée sous UV, déposée dans 350  $\mu$ l d'une solution Tris  
20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à  
65°C. Après dissolution de l'agarose, l'échantillon est  
extraît par un volume de phénol et centrifugé pendant 5  
min à 10 000 g. La phase aqueuse est reprise, extraite par  
35 un volume de phénol/chloroforme/alcool isoamylique  
(rapport 50/48/2) comme décrit dans préparation II1.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments  
5 HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 0,8%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment D BamHI/XbaI à 50 ng/µl.

Le plasmide utilisé pour cloner le fragment D  
10 est le vecteur pUC 18 commercialisé par Promega (TM). Ce plasmide comporte un gène de résistance à l'ampicilline.

Une quantité de 200 ng du plasmide pUC 18 est soumise à une digestion par les endonucléases de restriction BamHI et XbaI. Pendant 2 heures à 37°C. Après  
15 digestion, 2 µl de tampon porteur (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés au mélange et analysé sur gel d'agarose Seaplaque GTG (FMC, Rockland) 0,8%. La bande d'ADN correspondant au pUC18 linéarisé par BamHI et XbaI est prélevée et déposée dans 350 µl d'une solution de Tris  
20 20mM pH 8, EDTA 1 mM pH 8. Cet échantillon est incubé pendant 5 min à 65°C, extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) comme décrit dans  
25 PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 10 ng/µl du plasmide pUC18  
30 BamHI/XbaI.

La ligation du fragment D (BamHI/XbaI) au plasmide pUC18 (BamHI/XbaI) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries  
35 compétentes utilisées sont les DH5 alpha (Gibco BRL, Paris).

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), qui permet de purifier de l'ADN plasmidien super enroulé. Cette manipulation est réalisée suivant le protocole fourni par le fabricant. Environ 10 µg d'ADN plasmidien est purifié grâce à ce kit, lorsque le plasmide utilisé est le pUC18.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par les endonucléases de restriction BamHI et XbaI, afin de vérifier le clonage du fragment D. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment D est vérifiée sous UV.

Les réactions de séquence réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase (cf PREPARATION 1) permet de valider l'intégrité de la séquence du fragment D.

#### 5/ fragment E

Le fragment E est compris entre les sites DraIII (CACCTGGTG, nucléotides 1709 à 1717) et BamHI (GGATCC, nucléotides 2803 à 2808) présents dans le clone ADNc alpha22. Ce fragment code pour les acides aminés 567 à 936 compris dans le domaine hélicoïdal central de la chaîne proα1(I) humaine.

#### 6/ fragment F

Le fragment F est compris entre les sites BamHI (GGATCC, nucléotides 2803 à 2808) et EcoRI (GAATTC, nucléotides 4357 à 4362) présents dans le clone ADNc

alpha22. Ce fragment code pour les acides aminés 936 à 1192 compris dans le domaine hélicoïdal central et 1193 à 1454 du domaine carboxy-propeptide de la chaîne pro $\alpha$ 1(I) humaine.

5                    7/ Fragment G

Le fragment G comporte la séquence (nucléotides 4040 à 4392) codant pour la partie carboxy-terminale du domaine carboxy-propeptide (acides aminés 1346 à 1464) de la chaîne pro $\alpha$ 1(I) humaine. Ce fragment comporte aussi les  
10    nucléotides TAA (codon stop, bases 4393 à 4395) couplés à un site de restriction HinIII (AAGCTT).

Amplification du fragment G

Deux oligonucléotides de synthèse ont été  
utilisés.

15                    Oligonucléotide sens, BIOC25,  
                     5'- TATCTGCAGATGTGGCCATCCAGCTGACCT-3'-3'

Cet oligonucléotide comporte en 5' (nucléotides soulignés) la séquence permettant la création du site de restriction PstI (CTGCAG) qui sera utilisé comme site de  
20    clonage. A la suite de ce site, sont inclus les nucléotides 4040 à 4060 spécifiques de la séquence codante de la chaîne pro $\alpha$ 1(I) humaine.

                     Oligonucléotide antisens, BIOC23,  
                     5'- TATAAGCTTACAGGAAGCAGACAGGGCCAA-3'

25                    Cet oligonucléotide, présente en 5' (nucléotides soulignés) la séquence permettant la création du site de restriction HindIII (AAGCTT). En gras, est indiquée la séquence complémentaire du codon stop TAA, suivi du brin complémentaire des nucléotides 4373 à 4392 de la séquence  
30    codant pour la chaîne pro $\alpha$ 1(I) humaine.

La matrice ADN utilisée est le clone ADNc alpha22 qui comporte la séquence codant pour les acides aminés 171 à 1454 de la chaîne pro $\alpha$ 1(I) humaine et environ 500 pb de la région 3' non-traduite.

Clonage du fragment G

Après amplification, la solution est extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2). Le culot d'ADN est lavé avec 1 ml  
5 d'éthanol 70, lyophilisé, et resuspendu dans 35 µl d'eau milli-Q stérile.

La digestion par PstI est réalisée pendant 2 heures à 37°C dans le volume final de 40 µl.

Suite à la digestion enzymatique, 60 µl de  
10 tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION II1.

Après centrifugation à 4°C pendant 10 min à 12  
15 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 35 µl d'eau milli-Q stérile. La digestion par Hind III est réalisée pendant 2 heures à 37°C dans un volume final de 20 µl.

Suite à la digestion enzymatique, 8 µl de tampon  
20 stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 2% à bas point de fusion Nu Sieve-GTG (FMC, Rockland)cf. PREPARATION II2.

La bande correspondant au fragment G est  
25 découpée sous UV, déposée dans 350 µl d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par  
30 un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION II1.

Après centrifugation à 4°C pendant 10 min à 12  
000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une  
35 fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a



permis d'estimer la concentration de la solution purifiée du fragment G HindIII/PstI à 30 ng/μl.

Le plasmide utilisé pour cloner le fragment G est le vecteur pBluescript II SK(+) commercialisé par  
5 Stratagene. Ce plasmide comporte un gène de résistance à l'ampicilline.

Une quantité de 200 ng du plasmide pBluescript II SK(+) est soumise à une digestion par l'endonucléase de restriction PstI., dans un volume final de 20 μl, pendant  
10 2 heures à 37°C. Après digestion, 80 μl de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) comme décrit dans  
préparation II1.

15 Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17 μl d'eau milli-Q stérile. La digestion par Hind III est réalisée pendant 2 heures à 37°C dans un volume final de 20 μl.

20 Après digestion, 4 μl de tampon porteur (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés au mélange et analysé sur gel d'agarose Seaplaque GTG (FMC, Rockland) 0,8% cf PREPARATION II4. La bande d'ADN correspondant au pBluescript II SK(+) linéarisé par  
25 HindIII et PstI est prélevée et déposée dans 350 μl d'une solution de Tris 20mM pH 8, EDTA 1 mM pH 8. Cet échantillon est incubé pendant 5 min à 65°C, extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de  
30 phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION II1.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 μl d'eau milli-Q stérile afin  
35 d'obtenir une solution à 10 ng/μl du plasmide pBluescript II SK(+) HindIII/PstI.

La ligation du fragment G (HindIII/PstI) au plasmide pBluescript II SK(+) (HindIII/PstI) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries  
5 compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces  
10 cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega  
15 TM) suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par les endonucléases de restriction HindIII et BamHI, afin de vérifier le clonage du fragment G.

20 Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment G est vérifiée sous UV.

Les réactions de séquence réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de  
25 terminaison de chaînes et la T7 DNA polymerase (cf PREPARATION 1) permettent de valider l'intégrité de la séquence du fragment G.

#### 8/ Fragment H

Le fragment H comporte la séquence (nucléotides  
30 4031 à 4383) codant pour la partie carboxy-terminale du domaine carboxy-propeptide (acides aminés 1343 à 1461) de la chaîne proα1(I) humaine. Ce fragment comporte aussi les nucléotides TAA (codon stop, bases 4384 à 4386) couplés à un site de restriction HindIII (AAGCTT). En amont du codon  
35 stop sont inclus les codons Lys, Asp, Glu, Leu, site de rétention dans le réticulum endoplasmique.

## b/ Amplification du fragment H

Deux oligonucléotides de synthèse ont été utilisés.

Oligonucléotide sens, BIOC25,

5 5'- TATCTGCAGATGTGGCCATCCAGCTGACCT-3'-3'

Cet oligonucléotide comporte en 5' (nucléotides soulignés) la séquence permettant la création du site de restriction PstI (CTGCAG) qui sera utilisé comme site de clonage. A la suite de ce site, sont inclus les  
10 nucléotides 4031 à 4051 spécifiques de la séquence codante de la chaîne pro $\alpha$ 1(I) humaine.

Oligonucléotide antisens, BIOCKDEL,

5'TATAAGCTTATAGCTCATCTTTCAGGAAGCAGACAGGGCCAA  
CGTCGAAGC-3'

15 Cet oligonucléotide présente en 5' (nucléotides soulignés) la séquence permettant la création du site de restriction HindIII (AAGCTT). En gras, est indiquée la séquence complémentaire du codon stop TAA et des acides aminés KDEL, suivi du brin complémentaire des nucléotides  
20 4353 à 4383 de la séquence codant pour la chaîne pro $\alpha$ 1(I) humaine.

La matrice ADN utilisée est le clone ADNC alpha22 qui comporte 110 pb de la partie 5' non traduite et les 1911 premières bases codant pour la chaîne pro $\alpha$ 1(I)  
25 humaine.

Clonage du fragment H

Après amplification, la solution est extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION II.1. Le  
30 culot d'ADN est lavé avec 1 ml d'éthanol 70, lyophilisé, et resuspendu dans 35  $\mu$ l d'eau milli-Q stérile.

La digestion par PstI est réalisée pendant 2 heures à 37°C dans un volume final de 40  $\mu$ l.

Suite à la digestion enzymatique, 60  $\mu$ l de  
35 tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la

solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 35 µl d'eau milli-Q stérile. La digestion par HIND III est réalisée pendant 2 heures à 37°C dans un volume final de 40 µl.

Suite à la digestion enzymatique, 8 µl de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 2% à bas point de fusion Nu Sieve-GTG (FMC, Rockland) comme décrit dans PREPARATION III.

La bande correspondant au fragment H est découpée sous UV, déposée dans 350 µl d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment H HindIII/PstI à 20 ng/µl.

Le plasmide utilisé pour cloner le fragment H est le vecteur pBluescript II SK(+) commercialisé par Stratagene TM. Ce plasmide comporte un gène de résistance à l'ampicilline.

Une quantité de 200 ng du plasmide pBluescript II SK(+) est soumise à une digestion par l'endonucléase de restriction PstI. Dans un volume de 20 µl pendant 2 heures à 37°C. Après digestion, 80 µl de tampon TE (tris 10 mM,

EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION II1.

5           Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17 µl d'eau milli-Q stérile. La digestion par HINDIII est réalisée pendant 2 heures à 37°C dans un volume final de 20 µl.

10           Après digestion, 4 µl de tampon porteur (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés au mélange et analysé sur gel d'agarose Seaplaque GTG (FMC, Rockland) 0,8% cf. PREPARATION II4. La bande d'ADN correspondant au pBluescript II SK(+) linéarisé par  
15 HindIII et PstI est prélevée et déposée dans 350 µl d'une solution de Tris 20mM pH 8, EDTA 1 mM pH 8. Cet échantillon est incubé pendant 5 min à 65°C, extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de  
20 phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION II1.

          Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin  
25 d'obtenir une solution à 10 ng/µl du plasmide pBluescript II SK(+) HindIII/PstI.

          La ligation du fragment H (HindIII/PstI) au plasmide pBluescript II SK(+) (HindIII/PstI) est réalisée pendant 4 heures à température ambiante.

30           Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

          Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-  
35 Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

5 Une fraction aliquotée (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par les endonucléases de restriction HindIII et BamHI, afin de vérifier le clonage du fragment H. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de  
10 bromure d'éthidium à 40 ng/ml, et la présence du fragment H est vérifiée sous UV.

La dernière étape est de vérifier l'intégrité de la séquence du fragment H, la technique d'amplification par PCR (polymerase chain reaction) pouvant entraîner des  
15 mutations. Les réactions de séquence réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase (cf PREPARATION 1) permettent de valider l'intégrité de la séquence du fragment G.

20

#### EXEMPLE 1

Construction de gènes chimériques codant pour la protéine recombinante du collagène humain et permettant une expression dans les feuilles de  
25 tabac et les graines de tabac ou de colza.

##### 1/Obtention des fragment FG et FH

Cette étape consiste à cloner le fragment BamHI/EcoRI de 1554 pb (nucléotides 2803 à 4357) du clone alpha22 entre les sites BamHI (du vecteur pBluescript II  
30 SK(+)) et EcoRI (fragment G ou H) du clone pBluescript II SK(+)-fragment G ou H. Il en résulte l'obtention des clones pBluescript II SK(+)-FG et pBluescript II SK(+)-FH qui comportent entre les sites BamHI et HindIII la séquence codant pour les acides aminés 936 à 1192 du  
35 domaine hélicoïdal central, la totalité du domaine carboxy-propeptide (acides aminés 1193 à 1464) de la

chaîne pro $\alpha$ 1(I) humaine et le codon stop TAA couplé au site de restriction HindIII. Dans le cas du clone pBluescript II SK(+)-fragment FH, 4 codons spécifiques de la séquence KDEL (Lys, Asp, Glu, Leu) sont insérés entre la séquence codant pour le domaine carboxy-propeptide et le codon stop TAA.

Clonage du fragment F (BamHI/EcoRI) avec la partie EcoRI/HindIII du fragment G ou H.

Environ 1  $\mu$ g du clone ADNc alpha22 sont digérés par l'endonucléase de restriction BamHI pendant 90 min à 37°C. Les conditions utilisées sont: 1  $\mu$ g alpha22, 2  $\mu$ l de tampon de digestion E (Promega TM), 1  $\mu$ l de l'endonucléase de restriction BamHI (12 u/ $\mu$ l; Promega TM), qsp 20  $\mu$ l H<sub>2</sub>O milli-Q. La digestion est réalisée pendant 2 heures à 37°C.

Suite à la digestion enzymatique, 80  $\mu$ l de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans préparation II1.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 35  $\mu$ l d'eau milli-Q stérile. La digestion par PstI est réalisée pendant 2 heures à 37°C dans un volume final de 40  $\mu$ l.

Suite à la digestion enzymatique, 8  $\mu$ l de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 1% à bas point de fusion Sea-plaque-GTG (FMC, Rockland) à un voltage constant de 60 volts. Le tampon d'électrophorèse est du TBE 0,5X (Tris borate 45 mM, EDTA (Acide éthylène-diamine-tétra-acétique) 1mM, pH 8). Après migration, le gel est coloré pendant 15 min dans une solution de TBE 0,5X contenant 40 ng/ml de bromure d'ethidium.

La bande correspondant au fragment BamHI/EcoRI de 1554 bp est découpée sous UV, déposée dans 350 µl d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 1%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment BamHI/EcoRI du clone ADNc alpha22 à 12 ng/µl.

Le plasmide utilisé pour cloner le fragment BamHI/EcoRI du clone ADNc alpha22 est le clone pBluescript II SK(+)-fragment G ou H digéré par les enzymes BamHI et EcoRI. Ce clone comporte un gène de résistance à l'ampicilline.

Une quantité de 500 ng du plasmide pBluescript II SK(+)-fragment G ou H est soumise à une digestion par l'endonucléase de restriction BamHI. Dans un volume de 20 µl, pendant 2 heures à 37°C. Après digestion, 80 µl de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) comme décrit dans PREPARATION II.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17 µl d'eau milli-Q stérile. La digestion par EGRI est réalisée pendant 2 heures à 37°C dans un volume final de 20 µl.

Après digestion, 4 µl de tampon porteur (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés au



mélange et analysé sur gel d'agarose Seaplaque GTG (FMC, Rockland) 0,8% cf. PREPARATION II4. La bande d'ADN correspondant au pBluescript II SK(+)-fragment G ou H digéré par BamHI et EcoRI est prélevée et déposée dans 350  
5 µl d'une solution de Tris 20mM pH 8, EDTA 1 mM pH 8. Cet échantillon est incubé pendant 5 min à 65°C, extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2)  
10 comme décrit dans PREPARATION II1.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 20 ng/µl du plasmide pBluescript  
15 II SK(+)-fragment G ou H (BamHI/EcoRI).

La ligation du fragment BamHI/EcoRI du clone ADNc alpha22 au clone pBluescript II SK(+)-fragment G ou H (BamHI/EcoRI) est réalisée pendant 4 heures à température ambiante.

20 Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces  
25 cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et  
30 le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par les endonucléases de restriction HindIII et BamHI, afin de  
35 vérifier le clonage du fragment BamHI/EcoRI du clone ADNc alpha22 au clone pBluescript II SK(+)-fragment G ou H

(BamHI/EcoRI). Le site de restriction HindIII est couplé au codon stop TAA en 3' du site EcoRI.

Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment BamHI/HindIII de 1596 bp dans le cas du fragment FG et de 1608 pb dans le cas du fragment FH est vérifiée sous UV.

Les réactions de séquence ont été réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase permettent de valider l'intégrité des clones pBluescript II SK(+)-fragment FG et pBluescript II SK(+)-fragment FH.

#### 2/Obtention du fragment DE

Cette étape consiste à cloner le fragment DraIII/BamHI (nucléotides 1714 à 2803) de 1089 pb du clone alpha22 entre les sites DraIII (du fragment D) et BamHI (du plasmide pUC18) du clone pUC18-fragment D. Il en résulte l'obtention du clone pUC18-DE qui comporte entre les sites XbaI et BamHI la séquence codant pour les acides aminés 160 à 179 du domaine amino-télopeptide, les acides aminés 180 à 936 de la région hélicoïdale centrale de la chaîne pro $\alpha$ 1(I) humaine.

Clonage du fragment E (DraIII/BamHI) avec la partie DraIII/BamHI du fragment D.

Environ 1  $\mu$ g du clone ADNc alpha22 sont digérés par l'endonucléase de restriction BamHI pendant 90 min à 37°C.

Suite à la digestion enzymatique, 80  $\mu$ l de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) et centrifugée pendant 2 min à 10 000 g. La phase aqueuse est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration

finale en NaCl de 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 35 µl d'eau milli-Q stérile. La digestion par DraIII est réalisée pendant 2 heures à 37°C dans un volume final de 40 µl.

Suite à la digestion enzymatique, 8 µl de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 1% à bas point de fusion Sea-plaque-GTG (FMC, Rockland) cf EXEMPLE I1.

La bande correspondant au fragment DraIII/BamHI de 1089 pb du clone ADNc alpha22 est découpée sous UV, déposée dans 350 µl d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf PREPARATION III1.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment DraIII/BamHI du clone ADNc alpha22 à 20 ng/µl.

Le plasmide utilisé pour cloner le fragment DraIII/BamHI du clone ADNc alpha22 est le clone pUC18-fragment D digéré par les enzymes DraIII et BamHI. Ce clone comporte un gène de résistance à l'ampicilline.

Une quantité de 500 ng du plasmide pUC18-fragment D est soumise à une digestion par l'endonucléase de restriction BamHI. Dans un volume de 20 µl pendant 2 heures à 37°C. Après digestion, 80 µl de tampon TE (tris

10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf. PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17 µl d'eau milli-Q stérile. La digestion par DraIII est réalisée pendant 2 heures à 37°C dans un volume final de 20 µl.

Après digestion, 4 µl de tampon porteur (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés au mélange et analysé sur gel d'agarose Seaplaque GTG (FMC, Rockland) 0,8%. La bande d'ADN correspondant au pUC18-fragment D digéré par DraIII et BamHI est prélevée et déposée dans 350 µl d'une solution de Tris 20mM pH 8, EDTA 1 mM pH 8. Cet échantillon est incubé pendant 5 min à 65°C, extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 20 ng/µl du plasmide pUC18-fragment D (DraIII/BamHI).

La ligation du fragment DraIII/BamHI du clone ADNc alpha22 au clone pUC18-fragment D (DraIII/BamHI est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et

le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par les endonucléases de restriction XbaI et BamHI, afin de  
5 vérifier le clonage du fragment DraIII/BamHI du clone ADNc alpha22 au clone pUC18-fragment D (DraIII/BamHI). Le site de restriction XbaI est situé en 5' du clone pUC18-fragment D. Après électrophorèse, le gel est coloré  
10 pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment BamHI/XbaI de 2336 pb est vérifiée sous UV.

Les réactions de séquence ont été réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la  
15 technique de terminaison de chaînes et la T7 DNA polymerase permettent de valider l'intégrité du clone pUC18-fragment DE.

### 3/Obtention des fragments DEFG et DEFH

Cette étape consiste à cloner le fragment  
20 XbaI/BamHI (nucléotides 474 à 2803) de 2329 pb du clone pUC18-fragment DE entre les sites XbaI (du plasmide pBluescript II SK(+)) et BamHI (du fragment FG ou FH) du clone pBluescript II SK(+)-fragment FG ou FH. Il en résulte l'obtention des clones pBluescript II SK(+)-  
25 fragment DEFG et pBluescript II SK(+)-fragment DEFH qui comporte entre les sites XbaI et HindIII la séquence codant pour les acides aminés 160 à 179 du domaine amino-télopeptide, la totalité de la région hélicoïdale centrale (acides aminés 180 à 1192), la totalité du domaine  
30 carboxy-propeptide (acides aminés 1193 à 1464 de la chaîne proα1(I) humaine, et le codon stop TAA couplé au site de restriction HindIII. Dans le cas du clone pBluescript II SK(+)-fragment DEFH, 4 codons spécifiques de la séquence KDEL (Lys, Asp, Glu, Leu) sont insérés entre la séquence  
35 codant pour le domaine carboxy-propeptide et le codon stop TAA.

Clonage du fragment DE (XbaI/BamHI) avec la partie BamHI/HindIII du fragment FG ou FH.

Environ 1 µg du clone pUC18-fragment DE sont digérés par l'endonucléase de restriction BamHI pendant  
5 90 min à 37°C.

Suite à la digestion enzymatique, 80 µl de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2)  
10 cf PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 35 µl d'eau milli-Q stérile. La digestion par XbaI est réalisée pendant 2 heures à 37°C  
15 dans un volume final de 20 µl.

Suite à la digestion enzymatique, 8 µl de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 1% à bas point de fusion Sea-plaque-GTG  
20 (FMC, Rockland) EXEMPLE II.

La bande correspondant au fragment XbaI/BamHI de 2329 pb est découpée sous UV, déposée dans 350 µl d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose,  
25 l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf. PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés  
35 sur gel d'agarose type II (Sigma TM) 1%, ce qui nous a permis d'estimer la concentration de la solution purifiée

du fragment XbaI/BamHI du clone pUC18-fragment DE à 20 ng/µl.

Le plasmide utilisé pour cloner le fragment XbaI/BamHI du clone puC18-fragment DE est le clone  
5 pBluescript II SK(+)-fragment FG ou FH digéré par les enzymes XbaI et BamHI. Ce clone comporte un gène de résistance à l'ampicilline.

Une quantité de 500 ng du plasmide pBluescript II SK(+)-fragment FG ou FH est soumise à une digestion par  
10 l'endonucléase de restriction BamHI. Pendant 2 heures à 37°C. Après digestion, 80 µl de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) cf. PREPARATION III.

15 Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17 µl d'eau milli-Q stérile. La digestion par baI est réalisée pendant 2 heures à 37°C.

Après digestion, 4 µl de tampon porteur (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés au  
20 mélange et analysé sur gel d'agarose Seaplaque GTG (FMC, Rockland) 0,8%. La bande d'ADN correspondant au pBluescript II SK(+)-fragment FG ou FH digéré par XbaI et BamHI est prélevée et déposée dans 350 µl d'une solution  
25 de Tris 20mM pH 8, EDTA 1 mM pH 8. Cet échantillon est incubé pendant 5 min à 65°C, extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2)  
30 cf PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 20 ng/µl du plasmide pBluescript  
35 II SK(+)-fragment FG ou FH (XbaI/BamHI).

La ligation du fragment XbaI/BamHI du clone pUC18-fragment DE au clone pBluescript II SK(+)-fragment

FG ou FH (XbaI/BamHI) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par les endonucléases de restriction XbaI et HindIII, afin de vérifier le clonage du fragment XbaI/BamHI du clone pUC18-fragment DE au clone pBluescript II SK(+)-fragment FG ou FH (BamHI/HindIII). Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment XbaI/HindIII d'environ 3900 bp est vérifiée sous UV.

Les réactions de séquence ont été réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase permettent de valider l'intégrité des clones pBluescript II SK(+)-fragment DEFG et pBluescript II SK(+)-fragment DEFH.

4/Obtention des fragments ADEFG et ADEFH, clones ADNc complets spécifiques de la chaîne proα1(I) humaine.

Cette étape consiste à cloner le fragment SacII/NheI (nucléotides -4 à 479) de 495 pb du clone pBluescript II SK(+)-fragment A entre les sites SacII (du plasmide pBluescript II SK(+)) et NheI (du fragment DEFG ou DEFH) du clone pBluescript II SK(+)-fragment DEFG ou



DEFH. Il en résulte l'obtention des clones pBluescript II SK(+)-fragment ADEFG et pBluescript II SK(+)-fragment ADEFH qui comportent entre les sites SacII et HindIII la séquence codant pour totalité de la chaîne pro $\alpha$ 1(I) humaine, et le codon stop TAA couplé au site de restriction HindIII. Un autre site HindIII est localisé juste après le site SacII, ce qui permettra de sortir les fragments ADEFG et ADEFH par simple digestion par l'endonucléase de restriction HindIII. Dans le cas du clone pBluescript II SK(+)-fragment ADEFH, 4 codons spécifiques de la séquence KDEL (Lys, Asp, Glu, Leu) sont insérés entre la séquence codant pour le domaine carboxy-propeptide et le codon stop TAA. Le site NheI entraîne la substitution des acides aminés Asn et Phe par une lysine et une leucine en position 158 et 159 de la chaîne pro $\alpha$ 1(I) humaine.

Clonage du fragment A (SacII/NheI) avec la partie NheI/HindIII du fragment DEFG ou DEFH.

Environ 1  $\mu$ g du clone pBluescript II SK(+)-fragment A sont digérés par l'endonucléase de restriction SacII pendant 90 min à 37°C.

Suite à la digestion enzymatique, 80  $\mu$ l de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) cf. PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 35  $\mu$ l d'eau milli-Q stérile. A cette solution d'ADN sont ajoutés 4  $\mu$ l de tampon de digestion B (Promega TM), 1  $\mu$ l de l'endonucléase de restriction NheI (12 u/ $\mu$ l; Promega TM). La digestion est réalisée pendant 2 heures à 37°C.

Suite à la digestion enzymatique, 8  $\mu$ l de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse

sur gel d'agarose 2% à bas point de fusion NueSieve-GTG (FMC, Rockland) à un voltage constant de 60 volts. Le tampon d'électrophorèse est du TBE 0,5X (Tris borate 45 mM, EDTA (Acide éthylène-diamine-tétra-acétique) 1mM, pH 8).  
5 Après migration, le gel est coloré pendant 15 min dans une solution de TBE 0,5X contenant 40 ng/ml de bromure d'ethidium.

La bande correspondant au fragment SacII/NheI de 495 bp est découpée sous UV, déposée dans 350 µl d'une  
10 solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise,  
extraite par un volume de  
15 phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf. PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une  
20 fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment SacII/NheI du clone pBluescript II SK(+)-  
25 fragment A à 5 ng/µl.

Le plasmide utilisé pour cloner le fragment SacII/NheI du clone pBluescript II SK(+)-fragment A est le clone pBluescript II SK(+)-fragment DEFG ou DEFH digéré par les enzymes SacII et NheI. Ce clone comporte un gène  
30 de résistance à l'ampicilline.

Une quantité de 500 ng du plasmide pBluescript II SK(+)-fragment DEFG ou DEFH est soumise à une digestion par l'endonucléase de restriction SacII. Dans un volume de 20 µl pendant 2 heures à 37°C. Après digestion, 80 µl de  
35 tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de

phénol/chloroforme/ alcool isoamylique (rapport 50/48/2)  
cf PREPARATION II1.

Après centrifugation à 4°C pendant 10 min à 12  
000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70,  
5 lyophilisé et repris dans 17 µl d'eau milli-Q stérile. La  
digestion par NhcI est réalisée pendant 2 heures à 37°C.

Après digestion, 4 µl de tampon porteur (30%  
sucrose, 0,25% bleu de bromophénol) sont ajoutés au  
mélange et analysé sur gel d'agarose Seaplaque GTG (FMC,  
10 Rockland) 0,8% cf. PREPARATION II4. La bande d'ADN  
correspondant au pBluescript II SK(+)-fragment DEFG ou  
DEFH digéré par SacII et NheI est prélevée et déposée dans  
350 µl d'une solution de Tris 20mM pH 8, EDTA 1 mM pH 8.  
Cet échantillon est incubé pendant 5 min à 65°C, extrait  
15 par un volume de phénol et centrifugé pendant 5 min à 10  
000 g. La phase aqueuse est reprise, extraite par un  
volume de phénol/chloroforme/ alcool isoamylique (rapport  
50/48/2) cf. PREPARATION II1.

Après centrifugation à 4°C pendant 10 min à 12  
20 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70,  
lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin  
d'obtenir une solution à 20 ng/µl du plasmide pBluescript  
II SK(+)-fragment DEFG ou DEFH (SacII/NheI).

La ligation du fragment SacII/NheI du clone  
25 pBluescript II SK(+)-fragment A au clone pBluescript II  
SK(+)-fragment DEFG ou DEFH (SacII/NheI) pendant 4 heures  
à température ambiante.

Pour la transformation, les bactéries  
compétentes utilisées sont les DH5alpha (Gibco BRL,  
30 Paris).

Plusieurs colonies blanches sont testées. Pour  
cela, elles sont individuellement déposées dans 3 ml de L-  
Broth supplémenté en ampicilline (50 µg/ml), et ces  
cultures sont incubées sous agitation pendant une nuit à  
35 37°C.

Pour purifier l'ADN plasmidien, la culture est  
tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et

le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par l'endonucléase de restriction HindIII, afin de vérifier le clonage du fragment SacII/NheI du clone pBluescript II SK(+)-fragment A au clone pBluescript II SK(+)-fragment DEFG ou DEFH (NheI/HindIII). L'apparition d'un fragment HindIII de 4400 pb valide ces clonages. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment HindIII/HindIII de 4400 bp est vérifiée sous UV.

Les réactions de séquence ont été réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase permettent de valider l'intégrité des clones pBluescript II SK(+)-fragment ADEFG et pBluescript II SK(+)-fragment ADEFH.

Pour rappel, le clone pBluescript II SK(+)-fragment ADEFG comprend entre ces deux sites HindIII, la totalité de la séquence codant pour la chaîne pro $\alpha$ 1(I) humaine.

Le clone pBluescript II SK(+)-fragment ADEFH comprend entre ces deux sites HindIII, la totalité de la séquence codant pour la chaîne pro $\alpha$ 1(I) humaine. De plus, il génère en partie carboxy-terminale de cette chaîne, la séquence de rétention dans le réticulum endoplasmique Lys-Asp-Glu-Leu.

5/Obtention du fragment BDEFG, clone ADNC codant pour le signal peptide PRS de plante et la totalité de chaîne pro $\alpha$ 1(I) humaine à l'exception du domaine amino-propeptide.

Cette étape consiste à cloner le fragment SacII/NheI (nucléotides -3 à 75) de 90 pb du clone pBluescript II SK(+)-fragment B entre les sites SacII (du

plasmide pBluescript II SK(+)) et NheI (du fragment DEFG) du clone pBluescript II SK(+)-fragment DEFG. Il en résulte l'obtention du clone pBluescript II SK(+)-fragment BDEFG qui comporte entre les sites SacII et HindIII la séquence  
5 codant pour le signal peptide PRS de plante, la totalité de la chaîne pro $\alpha$ 1(I) humaine à l'exception du domaine amino-propeptide, et le codon stop TAA couplé au site de restriction HindIII. Un autre site HindIII est localisé juste après le site SacII, ce qui permettra de sortir le  
10 fragment BDEFG par simple digestion par l'endonucléase de restriction HindIII. Le site NheI entraîne la substitution des acides aminés Glu et Gly par une leucine et une Alanine en position 25 et 26 de la chaîne pro $\alpha$ 1(I) humaine.

15 Le signal peptide PRS (Pathogenesis-Related Protein S) de plante remplace le signal peptide propre à la chaîne pro $\alpha$ 1(I) humaine.

Clonage du fragment B (SacII/NheI) avec la partie NheI/HindIII du fragment DEFG.

20 Environ 2  $\mu$ g du clone pBluescript II SK(+)-fragment B sont digérés par l'endonucléase de restriction SacII pendant 90 min à 37°C.

Suite à la digestion enzymatique, 80  $\mu$ l de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la  
25 solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) cf. PREPARATION II.1.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70,  
30 lyophilisé et repris dans 34  $\mu$ l d'eau milli-Q stérile. La digestion par NheI est réalisée pendant 2 heures à 37°C.

Suite à la digestion enzymatique, 8  $\mu$ l de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse

sur gel d'agarose 2% à bas point de fusion NueSieve-GTG (FMC, Rockland) cf. PREPARATION III1.

La bande correspondant au fragment SacII/NheI de 90 bp est découpée sous UV, déposée dans 350 µl d'une  
5 solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise,  
extraite par un volume de  
10 phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf. PREPARATION III1.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une  
15 fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment SacII/NheI du clone pBluescript II SK(+)-  
20 fragment B à 1 ng/µl.

Le plasmide utilisé pour cloner le fragment SacII/NheI du clone pBluescript II SK(+)-fragment B est le clone pBluescript II SK(+)-fragment DEFG digéré par les enzymes SacII et NheI. Ce clone comporte un gène de  
25 résistance à l'ampicilline.

Une quantité de 500 ng du plasmide pBluescript II SK(+)-fragment DEFG est soumise à une digestion par l'endonucléase de restriction SacII. Dans un volume de 20 µl pendant 2 heures à 37°C. Après digestion, 80 µl de  
30 tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) cf. PREPARATION III1.

Après centrifugation à 4°C pendant 10 min à 12  
35 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17 µl d'eau milli-Q stérile. La digestion par NheI est réalisée pendant 2 heures à 37°C.

Après digestion, 4 µl de tampon porteur (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés au mélange et analysé sur gel d'agarose Seaplaque GTG (FMC, Rockland) 0,8% cf. PREPARATION II4. La bande d'ADN  
5 correspondant au pBluescript II SK(+)-fragment DEFG digéré par SacII et NheI est prélevée et déposée dans 350 µl d'une solution de Tris 20mM pH 8, EDTA 1 mM pH 8. Cet échantillon est incubé pendant 5 min à 65°C, extrait par un volume de phénol et centrifugé pendant 5 min à 10 000  
10 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) cf. PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70,  
15 lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 20 ng/µl du plasmide pBluescript II SK(+)-fragment DEFG (SacII/NheI).

La ligation du fragment SacII/NheI du clone pBluescript II SK(+)-fragment B au clone pBluescript II  
20 SK(+)-fragment DEFG (SacII/NheI) est pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

25 Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

30 Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus)  
35 est soumise à une digestion enzymatique par l'endonucléase de restriction HindIII, afin de vérifier le clonage du fragment SacII/NheI du clone pBluescript II SK(+)-fragment

B au clone pBluescript II SK(+)-fragment DEFG (NheI/HindIII). L'apparition d'un fragment HindIII de 4000 pb valide ces clonages. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment HindIII/HindIII de 4000 bp est vérifiée sous UV.

Les réactions de séquence ont été réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase permettent de valider l'intégrité du clone pBluescript II SK(+)-fragment BDEFG.

Pour rappel, le clone pBluescript II SK(+)-fragment BDEFG comprend entre ces deux sites HindIII, la séquence codant pour le signal peptide PRS de plante, et la totalité de la chaîne pro $\alpha$ 1(I) humaine, à l'exception du domaine amino-propeptide.

La création du site NheI, entraîne la substitution des acides aminés Glu et Gly par une leucine et une alanine en position 25 et 26 de la chaîne pro $\alpha$ 1(I) humaine.

6/Obtention du fragment BCDEFG, clone ADNc codant pour le signal peptide PRS de plante et la totalité de chaîne pro $\alpha$ 1(I) humaine.

Cette étape consiste à cloner le fragment NheI/NheI (nucléotides 72 à 479) de 402 pb du clone pBluescript II SK(+)-fragment C dans le site NheI (du fragment BDEFG) du clone pBluescript II SK(+)-fragment BDEFG. Il en résulte l'obtention du clone pBluescript II SK(+)-fragment BCDEFG qui comporte entre les sites SacII et HindIII la séquence codant pour le signal peptide PRS de plante, la totalité de la chaîne pro $\alpha$ 1(I) humaine, et le codon stop TAA couplé au site de restriction HindIII. Un autre site HindIII est localisé juste après le site SacII, ce qui permettra de sortir le fragment BCDEFG par simple digestion par l'endonucléase de restriction



HindIII. Les sites NheI entraîne la substitution des acides aminés Glu et Gly par une leucine et une alanine en position 25 et 26 de la chaîne pro $\alpha$ 1(I) humaine, et la substitution des acides aminés Asn et Phe par une lysine et une leucine en position 158 et 159 de la chaîne pro $\alpha$ 1(I) humaine.

Le signal peptide PRS de plante remplace le signal peptide propre à la chaîne pro $\alpha$ 1(I) humaine.

Clonage du fragment C (NheI/NheI) dans le site NheI du fragment BDEFG.

Environ 1  $\mu$ g du clone pBluescript II SK(+)-fragment C sont digérés par l'endonucléase de restriction NheI pendant 90 min à 37°C.

Suite à la digestion enzymatique, 4  $\mu$ l de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 2% à bas point de fusion NueSieve-GTG (FMC, Rockland) cf. PREPARATION III.

La bande correspondant au fragment NheI/NheI de 402 bp est découpée sous UV, déposée dans 350  $\mu$ l d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf. PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20  $\mu$ l d'eau milli-Q stérile. Une fraction aliquote de 2  $\mu$ l, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 2%, ce qui nous a permis d'estimer la concentration de la solution purifiée

du fragment NheI/NheI du clone pBluescript II SK(+)-fragment C à 5 ng/ $\mu$ l.

Le plasmide utilisé pour cloner le fragment NheI/NheI du clone pBluescript II SK(+)-fragment C est le  
5 clone pBluescript II SK(+)-fragment BDEFG linéarisé par l'enzyme NheI. Ce clone comporte un gène de résistance à l'ampicilline.

Une quantité de 250 ng du plasmide pBluescript II SK(+)-fragment BDEFG est soumise à une digestion par  
10 l'endonucléase de restriction NheI. Dans un volume de 20  $\mu$ l, pendant 2 heures à 37°C. Après digestion, 80  $\mu$ l de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2)  
15 cf. PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17  $\mu$ l d'eau milli-Q stérile. A cette solution d'ADN sont ajoutés 2  $\mu$ l de tampon 10X de la  
20 phosphatase alcaline d'intestinc de veaux (CIP, Promega TM), 1  $\mu$ l de CIP (1 u/ $\mu$ l; Promega TM). La déphosphoration est réalisée pendant 1 heures à 37°C. Cette étape permettra d'éviter l'autoligation de ce clone pendant l'étape de ligation.

25 Après digestion, 80  $\mu$ l de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) cf. PREPARATION III.

Après centrifugation à 4°C pendant 10 min à 12  
30 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20  $\mu$ l d'eau milli-Q stérile afin d'obtenir une solution à 10 ng/ $\mu$ l du plasmide pBluescript II SK(+)-fragment BDEFG (NheI, déphosphorylé).

La ligation du fragment NheI/NheI du clone  
35 pBluescript II SK(+)-fragment C au clone pBluescript II SK(+)-fragment BDEFG (NheI, déphosphorylé) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

Plusieurs colonies blanches sont testées. Pour  
5 cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est  
10 tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par l'endonucléase  
15 de restriction NheI, afin de vérifier le clonage du fragment NheI/NheI du clone pBluescript II SK(+)-fragment C au clone pBluescript II SK(+)-fragment BDEFG (NheI). L'apparition d'un fragment NheI de 402 pb valide ces clonages. Après électrophorèse, le gel est coloré pendant  
20 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment NheI de 402 pb est vérifiée sous UV.

Les réactions de séquence ont été réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la  
25 technique de terminaison de chaînes et la T7 DNA polymerase permettent de valider l'intégrité du clone pBluescript II SK(+)-fragment BCDEFG.

Pour rappel, le clone pBluescript II SK(+)-fragment BCDEFG comprend entre ces deux sites HindIII, la  
30 séquence codant pour le signal peptide PRS de plante, et la totalité de la chaîne proα1(I) humaine.

La création des sites NheI, entraîne la substitution des acides aminés Glu et Gly par une leucine et une alanine en position 25 et 26 de la chaîne proα1(I)  
35 humaine, et la substitution des acides aminés Asn et Phe

par une lysine et une leucine en position 158 et 159 de la chaîne pro $\alpha$ 1(I) humaine.

7/ Obtention des 4 constructions pBIOC21-collagène humain recombinant.

5 Les constructions des différents plasmides via l'utilisation des techniques d'ADN recombinant dérivent de pBIOC4. Le plasmide binaire dérive de pGA442 (An, 1986). Le plasmide dérivant de pBIOC4 et contenant la cassette d'expression ''PD35S-T35S'' est le plasmide pBIOC21.

10 Les différents éléments permettant de reproduire ces constructions sont par exemple contenus dans la description de la demande de brevet WO9633277 qui est incorporée par référence.

L'obtention des quatre constructions pBIOC21-collagène humain recombinant a consisté à passer les 4 constructions du plasmide pBluescript II SK(+) au plasmide pBIOC21.

15 Le plasmide pBIOC21 est utilisé pour permettre l'expression du collagène recombinant dans les feuilles et les graines de tabac ou de colza. Il apporte les séquences régulatrices suivantes:

20 a) le promoteur constitutif double 35S (PD35S) du virus de la mosaïque du chou-fleur (CaMV). Il correspond à une duplication de la séquence activant la transcription située en amont de l'élément TATA du promoteur 35S naturel.

2) la séquence terminatrice de transcription, le terminateur polyA 35S qui correspond à la séquence 3' non codante de la séquence du virus à ADN bicaténaire circulaire du CaMV produisant le transcrit 35S.

30 Clonage des fragments (HindIII/HindIII) des constructions pBluescript, dans le site de clonage HindIII du vecteur pBIOC21.

Environ 1  $\mu$ g du clone pBluescript II SK(+)-fragment ADEFG ou ADEFH ou BDEFG ou BCDEFG sont digérés

par l'endonucléase de restriction HindIII pendant 90 min à 37°C.

Suite à la digestion enzymatique, 4 µl de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 0,8% à bas point de fusion SeaPlaque-GTG (FMC, Rockland) cf. PREPARATION II4.

La bande correspondant au fragment HindIII/HindIII de 4400 bp est découpée sous UV, déposée dans 350 µl d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf. PREPARATION III1.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega TM) sont analysés sur gel d'agarose type II (Sigma TM) 0,8%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment HindIII/HindIII du clone pBluescript II SK(+)-fragment ADEFG ou ADEFH ou BDEFG ou BCDEFG à 25 ng/µl.

Le plasmide utilisé pour cloner le fragment HindIII/HindIII du clone pBluescript II SK(+)-fragment ADEFG ou ADEFH ou BDEFG ou BCDEFG est le clone pBIOC21 linéarisé par l'enzyme HindIII. Ce clone comporte un gène de résistance à la tétracycline.

Une quantité de 1 µg du plasmide pBIOC21 est soumise à une digestion par l'endonucléase de restriction HindIII. Dans un volume de 20 µl pendant 2 heures à 37°C. Après digestion, 80 µl de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) cf. PREPARATION III1

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17 µl d'eau milli-Q stérile. A cette solution d'ADN sont ajoutés 2 µl de tampon 10X de la phosphatase alcaline d'intestins de veaux (CIP, Promega TM), 1 µl de CIP (1 u/µl; Promega TM). La déphosphorylation est réalisée pendant 1 heures à 37°C. Cette étape permettra d'éviter l'autoligation de ce clone pendant l'étape de ligation.

10 Après digestion, 80 µl de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) cf. PREPARATION II1

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 40 ng/µl du plasmide pBIOC21 (HindIII, déphosphorylé).

La ligation du fragment HindIII/HindIII du clone pBluescript II SK(+)-fragment ADEFG ou ADEFH ou BDEFG ou BCDEFG au clone pBIOC21 (HindIII, déphosphorylé) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, 25 Paris).

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en tétracycline (40 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 30 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega TM), suivant le protocole fourni par le fabricant. Environ 35 1-2 µg d'ADN plasmidien est purifié grâce à ce kit, lorsque le plasmide utilisé est le pBIOC21.

Une fraction aliquote (5 µl sur les 60 obtenus) est soumise à une digestion enzymatique par l'endonucléase de restriction KpnI, afin de vérifier l'orientation du clonage du fragment HindIII du clone pBluescript II SK(+)-  
5 fragment ADEFG ou ADEFH ou BCDEFG au clone pBIOC21 (HindIII). L'apparition d'un fragment KpnI de 950 pb est spécifique d'une orientation correcte. Ce fragment KpnI de 950 pb résulte d'un site de restriction KpnI en position 2.8 kb du vecteur pBIOC21 et d'un site KpnI en position  
10 137 de la séquence codant pour la chaîne pro $\alpha$ 1(I) humaine. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment KpnI de 950 bp est vérifiée sous UV.

Dans le cas de la construction pBIOC21-fragment  
15 BDEFG, nous avons utilisé l'enzyme NheI et comparé son mode de digestion avec le vecteur pBIOC21)-fragment BCDEFG digéré par NheI. La seule différence est la présence d'un fragment NheI/NheI surnuméraire dans le cas du vecteur pBIOC21-fragment BCDEFG.

20 Les réactions de séquence ont été réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymerase permettent de valider l'intégrité des 4 constructions pBIOC21 réalisées.

25 Les séquences des 4 constructions vont être maintenant indiquées. Dans chaque cas, le signal peptide PRS de plante ou de la chaîne pro $\alpha$ 1(I) humaine sont soulignés, les substitutions d'acides aminés dû à la création d'un ou de deux sites de restriction NheI sont  
30 indiquées en gras ainsi que la séquence de rétention dans le réticulum endoplasmique KDEL. Les sites de reconnaissance à l'amino- et la carboxy-proteinase sont indiqués en relief.

Les constructions pBIOC21-collagène ainsi  
35 obtenues (fig 5 et 6) seront nommées ci après:

-pBIOC704 pour la construction contenant le fragment ADEFG

-pBIOC705 pour la construction contenant le fragment ADEFH

5 -pBIOC706 pour la construction contenant le fragment BDEFG

-pBIOC707 pour la construction contenant le fragment BCDEFG

10 L'ADN plasmidique des plasmides binaires, pBIOC704, pBIOC705, pBIOC706 et pBIOC707, a été introduit par transformation directe dans la souche LBA4404 d'Agrobacterium Tumefaciens selon le procédé de Holsters (1978). La validité des clones retenus est vérifiée par digestion enzymatique de l'ADN plasmidique introduit.

15 Restauration des acides aminés Asn et Phe, en position 158 et 159 de la chaîne pro $\alpha$ 1(I) humaine, dans les clones ADEFG, ADEFH et BCDEFG.

L'ensemble de ces clones sont inclus dans le vecteur pBluescript II SK(+) (Stratagene, TM). L'opération est d'échanger la séquence comprise entre les sites KpnI (bases 137 à 142) et DraIII (bases 1709 à 1720) de ces constructions par la séquence KpnI-DraIII du clone ADNc 1alpha3. Cette opération entraîne la restauration des acides aminés 158 et 159 de la chaîne pro $\alpha$ 1(I) (Asn et Phe). Cette opération a été menée puisque Morikawa (1980) ont montré que le résidu 159 (Phe) localisé trois acides aminés en amont du site de clivage Pro-Gln semblait jouer un rôle important pour la coupure de la chaîne pro $\alpha$ 1(I) par l'aminoprotéinase.

30 1/ Elimination du site de restriction KpnI du vecteur pUC18

Le vecteur pUC18 a été utilisé puisqu'il ne possède pas de site de restriction DraIII interne, à la différence du vecteur pBluescript II SK (+). L'élimination du site de restriction KpnI permettra l'utilisation des

35



enzymes DraIII et KpnI, dans le vue de l'échange des fragments recombinants KpnI-DraIII des clones ADEFG ou ADEFH ou BCDEFG par le fragment sauvage KpnI-DraIII du clone ADNc 1alpha3.

5 Pour cela, 100 ng du vecteur pUC18 sont digérés par l'endonucléase de restriction KpnI (Promega, TM) en suivant les recommandations du fabricant. Le vecteur linéarisé en KpnI a été ensuite traité par la nucléase Mung Bean (Promega, TM) selon les recommandations du  
10 fabricant afin d'obtenir un vecteur pUC18 KpnI à extrémités franches. Les étapes de ligation, transformation et de séquençage des clones obtenus sont réalisées comme présentées dans la partie préparation II.

2/ Transfert des fragments ADEFG, ADEFH et  
15 BCDEFG dans le vecteur pUC18minus KpnI

Pour cela 200 ng des constructions pBluescript II SK(+)-ADEFG ou ADEFH ou BCDEFG sont digérées par l'endonucléase de restriction HindIII (Promega, TM) selon les recommandations du fabricant.

20 Suite à la digestion enzymatique, 4 µl de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 0,8% à bas point de fusion SeaPlaque-GTG (FMC, TM) à un voltage constant de 60 volts. Le tampon  
25 d'électrophorèse est du TBE 0,5X (Tris borate 45 mM, EDTA (Acide éthylène-diamine-tétra-acétique) 1mM, pH 8). Après migration, le gel est coloré pendant 15 min dans une solution de TBE 0,5X contenant 40 ng/ml de bromure d'éthidium.

30 La bande correspondant au fragment ADEFG ou ADEFH ou BCDEFG est découpée sous UV, déposée dans 350 µl d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et  
35 centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) et

WO 96/27202

centrifugée pendant 2 min à 10 000 g. La phase aqueuse est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en NaCl de 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega, TM) sont analysés sur gel d'agarose type II (Sigma, TM) 0,8%, ce qui nous a permis d'estimer la concentration de la solution purifiée du fragment D BamHI/XbaI à 5 ng/µl.

Le plasmide utilisé pour cloner le fragment HindIII du clone pBluescript II SK(+)-fragment ADEFG ou ADEFH ou BCDEFG est le clone pUC18minus KpnI linéarisé par l'enzyme HindIII. Ce clone comporte un gène de résistance à l'ampicilline.

Une quantité de 250 ng du plasmide pUC18minusKpnI est soumise à une digestion par l'endonucléase de restriction HindIII selon les recommandations du fabricant. Suite à cette digestion, l'ADN est traité par la phosphatase alcaline d'intestins de veaux (CIP, Promega, TM) selon les recommandations du fabricant. Cette étape permettra d'éviter l'autoligation de ce clone pendant l'étape de ligation.

Après digestion, 80 µl de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) et centrifugée pendant 2 min à 10 000 g. La phase aqueuse est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en NaCl de 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin d'obtenir une solution à 10 ng/µl du plasmide pUC18minusKpnI (HindIII, déphosphorylé).

La ligation du fragment HindIII du clone pBluescript II SK(+)-fragment ADEFG ou ADEFH ou BCDEFG au vecteur pUC18minusKpnI (HindIII, déphosphorylé) est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, TM).

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega, TM), qui permet de purifier de l'ADN plasmidien super enroulé. Cette manipulation est réalisée suivant le protocole fourni par le fabricant. Environ 10 µg d'ADN plasmidien est purifié grâce à ce kit, lorsque le plasmide utilisé est le pUC18.

Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par l'endonucléase de restriction HindIII, afin de vérifier le clonage du fragment HindIII du clone pBluescript II SK(+)-fragment ADEFG ou ADEFH ou BCDEFG au clone pUC18minusKpnI (HindIII). Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment HindIII de 4400 pb est vérifiée sous UV.

Les réactions de séquences réalisées avec le kit "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymérase (cf PREPARATION 1) permet de valider l'intégrité du clonage.

3/ Echange des fragments KpnI-DraIII recombinants des clones ADEFG, ADEFH et BCDEFG par le fragment KpnI-DraIII sauvage du clone ADNc 1alpha3.

Pour cela 500 ng du clone ADNc 1alpha3 sont  
5 digérées par les endonucléases de restriction KpnI (Promega, TM) et DraIII (Boehringer Mannheim) selon les recommandations des fabricants.

Suite à la digestion enzymatique, 4 µl de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés  
10 à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 0,8% à bas point de fusion SeaPlaque-GTG (FMC, Rockland, TM) à un voltage constant de 60 volts. Le tampon d'électrophorèse est du TBE 0,5X (Tris borate 45 mM, EDTA (Acide éthylène-diamine-tétra-acétique) 1mM, pH 8).  
15 Après migration, le gel est coloré pendant 15 min dans une solution de TBE 0,5X contenant 40 ng/ml de bromure d'éthidium.

La bande correspondant au fragment KpnI-DraIII est découpée sous UV, déposée dans 350 µl d'une solution  
20 Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique  
25 (rapport 50/48/2) et centrifugée pendant 2 min à 10 000 g. La phase aqueuse est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en NaCl de 0,2 M, et précipitée  
30 16 heures à -20°C après apport de deux volumes d'éthanol 100.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une  
35 fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega, TM) sont analysés sur gel d'agarose type II (Sigma) 0,8%, ce qui nous a

permis d'estimer la concentration de la solution purifiée du fragment KpnI-DraIII du clone ADNc 1alpha3 à 5 ng/ $\mu$ l.

Le plasmide utilisé pour cloner le fragment KpnI-DraIII du clone ADNc 1alpha3 est le clone pUC18minus KpnI-fragment ADEFG ou ADEFH ou BCDEFG digéré par les enzymes KpnI-DraIII. Ce clone comporte un gène de résistance à l'ampicilline.

Une quantité de 250 ng du plasmide pUC18minusKpnI-fragment ADEFG ou ADEFH ou BCDEFG est soumise à une digestion par les endonucléases de restriction KpnI et DraIII selon les recommandations du fabricant.

Suite à la digestion enzymatique, 4  $\mu$ l de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 0,8% à bas point de fusion SeaPlaque-GTG (FMC, TM) à un voltage constant de 60 volts. Le tampon d'électrophorèse est du TBE 0,5X (Tris borate 45 mM, EDTA (Acide éthylène-diamine-tétra-acétique) 1mM, pH 8). Après migration, le gel est coloré pendant 15 min dans une solution de TBE 0,5X contenant 40 ng/ml de bromure d'éthidium.

La bande correspondant au vecteur pUC18minusKpnI-fragment ADEFG ou ADEFH ou BCDEFG KpnI-DraIII est découpée sous UV, déposée dans 350  $\mu$ l d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de phénol/chloroforme/alcool isoamylique (rapport 50/48/2) et centrifugée pendant 2 min à 10 000 g. La phase aqueuse est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en NaCl de 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une fraction aliquote de 2 µl, ainsi que 500 ng des fragments  
5 HindIII-EcoRI du phage lambda (Promega, TM) sont analysés sur gel d'agarose type II (Sigma) 0,8%, ce qui nous a permis d'estimer la concentration de la solution purifiée du vecteur pUC18minusKpnI-fragment ADEFG ou ADEFH ou BCDEFG KpnI-DraIII à 8 ng/µl.

10 La ligation du fragment KpnI-DraIII du clone ADNc 1alpha3 au vecteur pUC18minusKpnI-fragment ADEFG ou ADEFH ou BCDEFG KpnI-DraIII est réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries  
15 compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

Plusieurs colonies blanches sont testées. Pour cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en ampicilline (50 µg/ml), et ces  
20 cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega,  
25 TM), qui permet de purifier de l'ADN plasmidien super enroulé. Cette manipulation est réalisée suivant le protocole fourni par le fabricant. Environ 10 µg d'ADN plasmidien est purifié grâce à ce kit, lorsque le plasmide utilisé est le pUC18.

30 Une fraction aliquote (2 µl sur les 60 obtenus) est soumise à une digestion enzymatique par l'endonucléase de restriction NheI, afin de vérifier le clonage du fragment KpnI-DraIII du clone ADNc 1alpha3 au clone pUC18minusKpnI-fragment ADEFG ou ADEFH ou BCDEFG KpnI-  
35 DraIII. L'échange du fragment KpnI-DraIII recombinant par le fragment sauvage entraîne l'élimination d'un site de restriction NheI. Après électrophorèse, le gel est coloré

pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et l'élimination du site de restriction NheI est vérifiée sous UV.

Les réactions de séquences réalisées avec le kit  
5 "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymérase (cf PREPARATION 1) permet de valider l'intégrité du clonage.

Les fragments obtenus sont ADEFGphe159, ADEFHphe159 et BCDEFGphe159.

10 4/ Obtention des constructions pBIOC21-collagène humain recombinant phe159

Cette étape consiste à passer les 3 constructions du plasmide pUC18minusKpnI au plasmide pBIOC21.

15 Le plasmide pBIOC21 est utilisé pour permettre l'expression du collagène recombinant dans les feuilles de tabac ou de colza. Il apporte les séquences régulatrices suivantes:

a) le promoteur constitutif double-35S (PD35S)  
20 du CaMV. Il correspond à une duplication de la séquence activant la transcription située en amont de l'élément TATA du promoteur 35S naturel.

2) la séquence terminatrice de transcription, le terminateur polyA 35S qui correspond à la séquence 3' non  
25 codante de la séquence du virus à ADN bicaténaire circulaire de la mosaïque du chou-fleur produisant le transcrit 35S.

Les constructions des différents plasmides via l'utilisation des techniques d'ADN recombinant dérivent de  
30 pBIOC4. Le plasmide binaire dérive de pGA442 (An, 1986). Le plasmide dérivant de pBIOC4 et contenant la cassette d'expression "PD35S-T35S" est le plasmide pBIOC21.

Les différents éléments permettant de reproduire ces constructions sont par exemple contenus dans la  
35 description de la demande de brevet WO9633277 qui est incorporée par référence.

Clonage des fragments (HindIII/HindIII) des constructions pUC18minusKpnI, dans le site de clonage HindIII du vecteur pBIOC21.

Environ 1 µg du clone pUC18minusKpnI-fragment  
5 ADEFGphe159 ADEFHphe159 ou BCDEFGphe159 sont digérés par l'endonucléase de restriction HindIII selon les recommandations du fabricant.

Suite à la digestion enzymatique, 4 µl de tampon stop (30% sucrose, 0,25% bleu de bromophénol) sont ajoutés  
10 à la solution d'ADN qui est analysée par électrophorèse sur gel d'agarose 0,8% à bas point de fusion SeaPlaque-GTG (FMC, Rockland) à un voltage constant de 60 volts. Le tampon d'électrophorèse est du TBE 0,5X (Tris borate 45 mM, EDTA (Acide éthylène-diamine-tétra-acétique) 1mM, pH 8).  
15 Après migration, le gel est coloré pendant 15 min dans une solution de TBE 0,5X contenant 40 ng/ml de bromure d'éthidium.

La bande correspondant au fragment HindIII/HindIII de 4400 bp est découpée sous UV, déposée  
20 dans 350 µl d'une solution Tris 20 mM pH 8, EDTA 1 mM pH 8 et incubée pendant 5 min à 65°C. Après dissolution de l'agarose, l'échantillon est extrait par un volume de phénol et centrifugé pendant 5 min à 10 000 g. La phase aqueuse est reprise, extraite par un volume de  
25 phénol/chloroforme/alcool isoamylique (rapport 50/48/2) et centrifugée pendant 2 min à 10 000 g. La phase aqueuse est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en  
30 NaCl de 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile. Une  
35 fraction aliquote de 2 µl, ainsi que 500 ng des fragments HindIII-EcoRI du phage lambda (Promega, TM) sont analysés sur gel d'agarose type II (SigmaTM) 0,8%, ce qui nous a



permis d'estimer la concentration de la solution purifiée du fragment HindIII/HindIII du clone pUC18minusKpnI-fragment ADEFGphe159 ou ADEFHphe159 ou BCDEFGphe159 à 25 ng/μl.

5 Le plasmide utilisé pour cloner le fragment HindIII/HindIII du clone pUC18minusKpnI-fragment ADEFGphe159 ou ADEFHphe159 ou BCDEFGphe159 est le clone pBIOC21 linéarisé par l'enzyme HindIII. Ce clone comporte un gène de résistance à la tétracycline.

10 Une quantité de 1 μg du plasmide pBIOC21 est soumise à une digestion par l'endonucléase de restriction HindIII. Cette digestion est réalisée dans un volume de 20 μl en présence de 2 μl de tampon B (Promega, TM), 1 μl d'enzyme HindIII (12 u/μl; Promega, TM), pendant 2 heures  
15 à 37°C. Après digestion, 80 μl de tampon TE (tris 10 mM, EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) et centrifugée pendant 2 min à 10 000 g. La phase aqueuse est extraite au  
20 chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en NaCl de 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

25 Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 17 μl d'eau milli-Q stérile. A cette solution d'ADN sont ajoutés 2 μl de tampon 10X de la phosphatase alcaline d'intestins de veaux (CIP, Promega,  
30 TM), 1 μl de CIP (1 u/μl; Promega, TM). La déphosphorylation est réalisée pendant 1 heures à 37°C. Cette étape permettra d'éviter l'autoligation de ce clone pendant l'étape de ligation.

Après digestion, 80 μl de tampon TE (tris 10 mM,  
35 EDTA 1 mM, pH 8) sont ajoutés à la solution d'ADN qui est extraite par un volume de phénol/chloroforme/ alcool isoamylique (rapport 50/48/2) et centrifugée pendant 2 min

à 10 000 g. La phase aqueuse est extraite au chloroforme/alcool isoamylique (rapport 24/1) et centrifugée quelques secondes à 10 000 g. La phase aqueuse est reprise, ajustée à une concentration finale en NaCl de 5 0,2 M, et précipitée 16 heures à -20°C après apport de deux volumes d'éthanol 100.

Après centrifugation à 4°C pendant 10 min à 12 000 g, le culot d'ADN est lavé par 1 ml d'éthanol 70, lyophilisé et repris dans 20 µl d'eau milli-Q stérile afin 10 d'obtenir une solution à 40 ng/µl du plasmide pBIOC21 (HindIII, déphosphorylé).

La ligation du fragment HindIII/HindIII du clone pUC18minusKpnI-fragment ADEFGphe159 ou ADEFHphe159 ou BCDEFGphe159 au clone pBIOC21 (HindIII, déphosphorylé) est 15 réalisée pendant 4 heures à température ambiante.

Pour la transformation, les bactéries compétentes utilisées sont les DH5alpha (Gibco BRL, Paris).

Plusieurs colonies blanches sont testées. Pour 20 cela, elles sont individuellement déposées dans 3 ml de L-Broth supplémenté en tétracycline (40 µg/ml), et ces cultures sont incubées sous agitation pendant une nuit à 37°C.

Pour purifier l'ADN plasmidien, la culture est 25 tout d'abord centrifugée à 1500 g pendant 10 min à 4°C, et le culot bactérien est traité avec le kit Wizard (Promega, TM), qui permet de purifier de l'ADN plasmidien super enroulé. Cette manipulation est réalisée suivant le protocole fourni par le fabricant. Environ 1-2 µg d'ADN 30 plasmidien est purifié grâce à ce kit, lorsque le plasmide utilisé est le pBIOC21.

Une fraction aliquote (5 µl sur les 60 obtenus) est soumise à une digestion enzymatique par l'endonucléase de restriction KpnI, afin de vérifier l'orientation du 35 clonage du fragment HindIII du clone pUC18minusKpnI-fragment ADEFGphe159 ou ADEFHphe159 ou BCDEFGphe159 au clone pBIOC21 (HindIII). L'apparition d'un fragment KpnI

de 950 pb est spécifique d'une orientation correcte. Ce fragment KpnI de 950 pb résulte d'un site de restriction KpnI en position 2.8 kb du vecteur pBIOC21 et d'un site KpnI en position 137 de la séquence codant pour la chaîne  
5 pro $\alpha$ 1(I) humaine. Après électrophorèse, le gel est coloré pendant 15 min dans une solution de bromure d'éthidium à 40 ng/ml, et la présence du fragment KpnI de 950 bp est vérifiée sous UV.

Les réactions de séquences réalisées avec le kit  
10 "T7 sequencing" (Pharmacia TM) qui utilise la technique de terminaison de chaînes et la T7 DNA polymérase (cf PREPARATION 1) permet de valider l'intégrité du clonage.

Les constructions pBIOC21-collagène ainsi obtenues seront nommées ci après:

- 15 -pBIOC708 pour la construction contenant le fragment ADEFGphe159
- pBIOC709 pour la construction contenant le fragment ADEFHphe159
- pBIOC710 pour la construction contenant le  
20 fragment BCDEFHphe159

## EXEMPLE 2

Construction de gènes chimériques codant pour la protéine recombinante du collagène humain  
25 et permettant une expression dans les semences de maïs.

L'expression constitutive dans les semences de maïs du cDNA codant la chaîne pro- $\alpha$ 1(I) du collagène humain a nécessité les séquences régulatrices suivantes :

- 30 1. le promoteur du gène de zéine de maïs (Pzéine) contenu dans le plasmide p63. Il permet une expression dans l'albumen des semences de maïs ;
- 2. la séquence terminatrice de transcription, terminateur polyA NOS, qui correspond à la région en 3'
- 35 non codante du gène de la nopaline synthase du plasmide Ti d'*Agrobacterium tumefaciens* souche à nopaline.

Les différents éléments permettant de reproduire ces constructions sont par exemple contenus dans la description de la demande de brevet WO9633277 qui est incorporé par référence.

5 A partir des plasmides, les fragments ADEFG, ADEFH, BDEFG, et BCDEFG ont été isolés par digestion par HindIII, purifiés par électrophorèse sur gel d'agarose 0,8%, électroélués, soumis à la précipitation alcoolique et séchés. Puis ces fragments d'ADN ont été traités par  
10 l'enzyme de Klenow (New England Biolabs TM) selon les recommandations du fabricant, soumis à l'extraction phénol-chloroforme, à la précipitation alcoolique et repris dans 10 µl H<sub>2</sub>O.

Le plasmide p63 a été doublement digéré par SacI  
15 et BamHI, purifié par électrophorèse sur gel d'agarose 0,8%, électroélué, soumis à la précipitation alcoolique et séché. Puis il a été traité par l'enzyme T4 polymérase (New England Biolabs TM) et déphosphorylé par l'enzyme phosphatase alcaline d'intestin de veau (Boehringer  
20 Mannheim) selon les recommandations des fabricants.

Chaque fragment ADEFG, ADEFH, BDEFG, et BCDEFG obtenu comme décrit ci-dessus, a été ligué au plasmide p63 traité comme décrit ci-dessus.

La ligation et la transformation des bactéries  
25 compétentes *Escherichia coli* souche DH5a ont été réalisées selon les procédés standard. Les plasmides obtenus sont appelés respectivement pBIOC700, pBIOC701, pBIOC702 et pBIOC703 et contiennent respectivement les fragments ADEFG, ADEFH, BDEFG et BCDEFG.

30 Les acides aminés Asn et Phe, en positions 158 et 159 de la chaîne proα 1(I) humaine, ont été restaurés dans les clones ADEFG, ADEFH et BCDEFG pour obtenir ADEFGphe159, ADEFHphe159 et BCDEFHphe159, respectivement. Cette modification est décrite dans l'exemple 1.

35 A partir des plasmides, les fragments ADEFGphe159, ADEFHphe159 et BCDEFHphe159 ont été isolés

par digestion par HindIII, purifiés par électrophorèse sur gel d'agarose 0,8%, électroélués, soumis à la précipitation alcoolique et repris dans 10µl H<sub>2</sub>O. Ces fragments ainsi préparés ont été ligués au plasmide p63  
5 traité comme décrit ci-dessus. Les plasmides obtenus sont appelés pBIOC711, pBIOC712 et pBIOC713 et contiennent les fragments ADEFGphe159, ADEFHphe159 et BCDEFHphe159, respectivement.

10                   **EXEMPLE 3**  
                    **Obtention de plants de colza**  
                    **transgéniques.**

Les graines de colza de printemps (*Brassica napus* cv WESTAR ou lignées Limagrain) sont désinfectées  
15 pendant 40 minutes dans une solution de Domestos TM à 15%. Après 4 rinçages à l'eau stérile, les graines sont mises à germer, à raison de 20 graines par pot de 7 cm de diamètre et 10 cm de haut, sur du milieu minéral de Murashige et Skoog (Sigma M 5519) avec 30g/l de saccharose et solidifié  
20 avec de l'agargel à 5g/l. Ces pots sont placés dans une chambre de culture à 26°C avec une photopériode de 16h/8h et sous une intensité lumineuse de l'ordre de 80 µE m<sup>-2</sup> s<sup>-1</sup>.

Après 5 jours de germination, les cotylédons  
25 sont prélevés stérilement en coupant chaque pétiole environ 1 mm au-dessus du noeud cotylédonaire.

Parallèlement une préculture d'*Agrobacterium tumefaciens* souche LBA4404, contenant les plasmides, est réalisée en erlenmeyer de 50 ml, pendant 36 h à 28°C dans  
30 10 ml de milieu bactérien 2YT complémenté avec les antibiotiques utiles à la sélection de la souche utilisée.

Cette préculture sert à ensemercer à 1% une nouvelle culture bactérienne effectuée dans les mêmes conditions. Au bout de 14 h la culture est centrifugée 15  
35 min à 3000 g et les bactéries sont reprises dans un volume équivalent de milieu de germination liquide. Cette

suspension est distribuée dans des boîtes de pétri de 5 cm de diamètre à raison de 5 ml/boîte.

L'extrémité sectionnée du pétiole est immergée quelques secondes dans la solution d'agrobactéries ainsi préparées, puis le pétiole est enfoncé de quelques millimètres dans le milieu de régénération. Ce milieu a la même composition de base que le milieu de germination avec en plus 4 mg/l de benzyl-amino-purine (BAP), phytohormone favorisant la néoformation de bourgeons. Dix explants (cotylédon avec pétiole) sont mis en culture par boîte de pétri de 9 cm de diamètre .

Après 2 jours de coculture dans les mêmes conditions environnementales que les germinations, les explants sont repiqués dans des boîtes phytatray (Sigma TM, référence P1552) contenant le milieu précédent complémenté avec un agent sélectif: 45 mg/l de sulfate de kanamycine et un bactériostatique: mélange de 1/6 (en poids) de sel de potassium d'acide clavulanique et de 5/6 sel de sodium d'amoxicilline (Augmentin TM injectable) à raison de 600 mg/l.

Deux fois de suite, à 3 semaines d'intervalle, les explants sont repiqués stérilement sur du milieu neuf dans les mêmes conditions.

Les bourgeons verts apparus à la fin du deuxième ou du troisième repiquage sont séparés de l'explant et mis en culture individuellement dans des pots transparents de 5 cm de diamètre et de 10 cm de haut contenant un milieu identique au précédent mais dépourvu de BAP. Après 3 semaines de culture la tige du bourgeon transformé est sectionnée et le bourgeon est repiqué dans un pot de milieu frais. Au bout de trois à quatre semaines les racines sont assez développées pour permettre l'acclimatation de la plantule dans un phytotron. Les bourgeons qui ne sont pas verts ou enracinés sont éliminés. Ces plantules sont alors transplantées dans des godets de 7 cm de côté remplis de terreau (norme NF U44551: 40% tourbe brune, 30% bruyère tamisée et 30%

sable) saturé en eau. Après deux semaines d'acclimatation en phytotron (température 21°C, photopériode 16h/8h et 84% d'humidité relative), les plantules sont repotées dans des pots de 12 cm de diamètre remplis du même terreau enrichi en engrais retard puis transportées en serre (classe S2) régulée à 18°C, avec deux arrosages quotidiens de 2 minutes à l'eau.

Dès l'apparition de fleurs celles-ci sont de façon à empêcher la fécondation croisée.

Lorsque les siliques sont arrivées à maturité, elles sont récoltées, séchées, puis battues. Les graines obtenues servent au dosage de l'activité biochimique. La sélection de la descendance transgénique se fait par germination sur un milieu contenant du sulfate de kanamycine à raison de 100 à 150 mg/l (selon les génotypes). Les conditions opératoires sont identiques à celles décrites ci-dessus à ceci près que les germinations sont effectuées en tubes de verre avec une seule graine par tube. Seules les plantules développant des racines secondaires les trois premières semaines sont acclimatées en phytotron avant d'être passées en serre.

#### EXEMPLE 4

##### Obtention de plants transgéniques de tabac

Les plants de tabac utilisés pour les expériences de transformation (*Nicotiana tabacum* var. Xanthi NC et PBD6) sont cultivées *in vitro* sur le milieu de base de Murashige et Skoog (1962) additionné des vitamines de Gamborg et al. (1968, Sigma référence M0404) de saccharose à 20 g/l et d'agar à 8 g/l. Le pH du milieu est ajusté à 5,8 avec une solution de potasse avant autoclavage à 120°C pendant 20 min. Les plantules de tabac sont repiquées par bouture des entre-noeuds tous les 30 jours sur ce milieu de multiplication MS20.

Toutes les cultures *in vitro* sont réalisées en enceinte climatisée, dans les conditions définies ci-dessous:

- intensité lumineuse de  $30 \mu\text{E.m}^{-2}.\text{s}^{-1}$ ;  
photopériode de 16h;

- thermopériode de 26°C le jour, 24°C la nuit.

La technique de transformation utilisée est  
5 dérivée de celle de Horsch et al. (1985).

Une préculture d'*Agrobacterium tumefaciens*  
souche LBA4404 contenant les plasmides est réalisée durant  
48h à 28°C sous agitation, dans du milieu LB additionné  
des antibiotiques adéquats (rifampicine et tétracycline).  
10 La préculture est ensuite diluée au 50ème dans le même  
milieu et cultivée dans les mêmes conditions. Après une  
nuit, la culture est centrifugée (10 min., 3000 g), les  
bactéries sont reprises dans un volume équivalent de  
milieu MS30 (30 g/l saccharose) liquide et cette  
15 suspension est diluée au 10ème.

Des explants d'environ 1 cm<sup>2</sup> sont découpés à  
partir des feuilles des plantules décrites ci-dessus. Ils  
sont ensuite mis au contact de la suspension bactérienne  
pendant 1h, puis séchés rapidement sur papier filtre et  
20 placés sur un milieu de coculture (MS30 solide).

Après 2 jours, les explants sont transférés en  
boîtes de Pétri sur le milieu de régénération MS30,  
contenant un agent sélectif, la kanamycine (200 mg/l), un  
bactériostatique, l'augmentin (400 mg/l) et les hormones  
25 nécessaires à l'induction de bourgeons (BAP, 1 mg/l et  
ANA, 0,1 mg/l). Un repiquage des explants est effectué sur  
le même milieu après 2 semaines de culture. Après 2  
semaines supplémentaires, les bourgeons sont repiqués en  
boîtes de Pétri sur le milieu de développement composé du  
30 milieu MS20 additionné de kanamycine et d'augmentin. Après  
15 jours, les bourgeons sont repiqués de moitié.  
L'enracinement prend environ 20 jours, au terme desquels  
les plantules peuvent être clonées par bouture d'entre-  
noeuds ou sorties en serre.



**EXEMPLE 5****Obtention de plantes de maïs transgéniques.**

a) Obtention et utilisation de cal de maïs comme  
5 cible pour la transformation génétique.

La transformation génétique du maïs, quelle que  
soit la méthode employée (électroporation; *Agrobacterium*,  
microfibres, canon à particules), requiert généralement  
l'utilisation de cellules indifférenciées en divisions  
10 rapides ayant conservé une aptitude à la régénération de  
plantes entières. Ce type de cellules compose le cal  
friable embryogène (dit de type II) de maïs.

Ces cals sont obtenus à partir d'embryons  
immatures de génotype H1 II ou (A188 x B73) selon la  
15 méthode et sur les milieux décrits par Armstrong (1994).  
Les cals ainsi obtenus sont multipliés et maintenus par  
repiquages successifs tous les quinze jours sur le milieu  
d'initiation.

Des plantules sont ensuite régénérées à partir  
20 de ces cals en modifiant l'équilibre hormonal et osmotique  
des cellules selon la méthode décrite par Vain et al  
(1989). Ces plantes sont ensuite acclimatées en serre où  
elles peuvent être croisées ou autofécondées.

b) Utilisation du canon à particules pour la  
25 transformation génétique du maïs.

Le paragraphe précédent décrit l'obtention et la  
régénération des lignées cellulaires nécessaires à la  
transformation ; on décrit ici une méthode de  
transformation génétique conduisant à l'intégration stable  
30 des gènes modifiés dans le génome de la plante. Cette  
méthode repose sur l'utilisation d'un canon à; les  
cellules cibles sont des fragments de cals décrits dans le  
paragraphe 1. Ces fragments d'une surface de 10 à 20 mm<sup>2</sup>  
ont été disposés, 4 h avant bombardement, à raison de 16  
35 fragments par boîte au centre d'une boîte de petri  
contenant un milieu de culture identique au milieu  
d'initiation, additionné de 0,2 M de mannitol + 0,2 M de

sorbitol. Les plasmides portant les gènes à introduire sont purifiés sur colonne Qiagen TM, en suivant les instructions du fabricant. Ils sont ensuite précipités sur des particules de tungsten (M10) en suivant le protocole  
5 décrit par Klein (Nature (1987) 327:70-73). Les particules ainsi enrobées sont projetées vers les cellules cibles à l'aide du canon .

Les boîtes de cals ainsi bombardées sont ensuite scellées , puis cultivées à l'obscurité à 27°C. Le premier  
10 repiquage a lieu 24 h après, puis tous les quinze jours pendant 3 mois sur milieu identique au milieu d'initiation additionné d'un agent sélectif dont la nature et la concentration peuvent varier selon le gène utilisé (voir paragraphe 3). Les agents sélectifs utilisables consistent  
15 généralement en composés actifs de certains herbicides (Basta, Round up) ou certains antibiotiques (Hygromycine, Kanamycine...).

On obtient après 3 mois ou parfois plus tôt, des cals dont la croissance n'est pas inhibée par l'agent  
20 sélection, habituellement et majoritairement composés de cellules résultant de la division d'une cellule ayant intégré dans son patrimoine génétique une ou plusieurs copies du gène de sélection. La fréquence d'obtention de tels cals est d'environ 0,8 cal par boîte bombardée.

25 Ces cals sont identifiés, individualisés, amplifiés puis cultivés de façon à régénérer des plantules (cf. paragraphe a). Afin d'éviter toute interférence avec des cellules non transformées toutes ces opérations sont menées sur des milieux de culture contenant l'agent  
30 sélectif.

Les plantes ainsi régénérées sont acclimatées puis cultivées en serre où elles peuvent être croisées ou autofécondées.

#### 35 EXEMPLE 6 : EXTRACTION ET CRIBLAGE

Extraction à partir de plants de tabac génétiquement transformés selon l'invention

45 plants numérotés de 1 à 22 lorsqu'ils ont été transfectés avec la construction pBIOC707 (comprenant le N-propeptide, la triple hélice majeure et le C-propeptide) de 23 à 45 lorsqu'ils ont été transfectés par la construction pBIOC706 (comprenant la triple hélice majeure et le C-propeptide) sont coupés, pesés et congelés dans l'azote liquide. Les différents plants sont alors broyés dans l'azote liquide à l'aide d'un homogénéiseur pendant 2 min afin d'obtenir une poudre fine puis stockés jusqu'à leur utilisation à -20°C.

#### **Extraction par l'acide acétique:**

La poudre obtenue pour chaque plant y compris les plants témoins non-transformés, est reprise par une solution d'acide acétique 0.5M contenant 200 mM NaCl et des inhibiteurs de protéases : 1mM phényl méthyl sulfonide (PMSF), 1 mM N-ethyl maléimide (NEM), 2.5 mM acide tétra-acétique diamine éthylène (EDTA), à raison de 1 gramme de poudre pour 4 ml de solution. Le mélange est agité à 4°C pendant 60 heures. Après centrifugation de chaque échantillon à 20 000 g (rotor 12148, Sigma TM) pendant 30 minutes à 4°C. Les culots sont lavés 2 fois en eau distillée et stockés à -20°C. Une fraction des surnageants (100 µl) est prélevée pour réaliser un dosage protéique sur l'extrait total. Le reste des surnageants est stockés à -20°C.

#### **Dosage protéique de l'extrait total:**

Les fractions prélevées 100µl pour l'extrait acido-soluble et 20µl pour l'extrait en tampon neutre sont dosés par l'utilisation d'une solution de Coomassie prête à l'emploi. Le dosage est réalisé en plaque multipuits et la lecture est faite sur un lecteur de plaques (SLT Lab Instruments, Austria) à une longueur d'onde de 620nm. Une gamme étalon est réalisée à partir d'une solution de BSA à 100µg/ml de 0 à 10 µg/puits: 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100µl sont déposés dans chaque puits et complétés lorsque nécessaire à 100µl final. Les différents échantillons sont également déposés dans les puits et

complétés lorsque nécessaire à 100µl dans le tampon d'extraction. Les références se font par 100µl d'eau distillée pour la gamme étalon et 100µl de tampon d'extraction pour les échantillons. . 100µl de réactif  
5 sont alors ajoutés dans chaque puits, la lecture est immédiate. Les valeurs obtenues pour chacune des références seront déduites respectivement des valeurs obtenues pour la gamme étalon et celles des échantillons. Une courbe étalon est tracée et la concentration en  
10 protéine des échantillons est déduite à partir de cette courbe.

#### Criblage des positifs:

Le criblage s'opère de deux façons :

1. Criblage des plants (après broyage et  
15 extraction) par gel d'électrophorèse suivi d'un immunotransfert. Le collagène témoin utilisé sera du collagène I extrait de tissus humain ou bovin par l'acide acétique et purifié par précipitation par les sels. La forme moléculaire du collagène I la plus répandue dans les  
20 tissus est l'hétérotrimère contenant deux chaînes  $\alpha 1$  et  $\alpha 2$ : [ $\alpha 1(I)$   $2\alpha 2(I)$ ]. Dans ce cas, la chaîne d'intérêt  $\alpha 1$  migre légèrement au dessus de 116KDa, elle ne contient que la triple hélice majeure et les télépeptides, les N et C-propeptides ont été clivés lors de la maturation du  
25 collagène et ne sont donc plus présents.

2. Criblage des positifs par immunofluorescence indirecte sur coupes congelées des différents plants.

Dans les 2 cas, l'anticorps utilisé est un anticorps polyclonal fait chez le lapin spécifique du  
30 collagène I, il reconnaît le collagène I humain et bovin et indifféremment les chaînes  $\alpha 1(I)$  et  $\alpha 2(I)$  (référence 20121, Institut Pasteur, Lyon ); Un autre anticorps peut être également utilisé: il s'agit d'un anticorps monoclonal Sp1D8 (Hybridoma Bank, USA) dirigé contre le  
35 site de clivage du N-propeptide de type I. Ces deux

anticorps peuvent être utilisés sur du collagène I dénaturé (après gel d'électrophorèse et transfert des protéines) et sur du collagène natif (en immunofluorescence).

5           1. Criblage après analyse des plants par électrophorèse suivie d'un immunotransfert.

          Méthode analytique

          Le criblage est réalisé sur les extraits en acide acétique. Les surnageants et culots sont analysés.

10           Electrophorèse:

Les surnageants:

          500µl des surnageants, soit 15 à 25 µg de protéines totales, sont précipités par l'addition de 50 µl d'acide trichloroacétique 100% pendant 30 minutes à  
15 température ambiante. Les échantillons sont centrifugés à 20 000g pendant 15 minutes à 4°C, les culots sont lavés 2 fois en éthanol 100% froid. Les précipités sont alors repris par 20µl de tampon porteur (Tris-HCl 60mM pH 6,8, contenant 2% sodium dodecyl sulfate, 10% glycérol, 0.002%  
20 de bleu de bromophénol) en présence d'un agent réducteur le dithiothreitol (DTT) à 10 mM final, puis dénaturés pendant 3 minutes à 100°C.

Les culots:

          Les culots, lavés en eau distillée pour éviter  
25 la présence d'acide et de sels provenant du tampon d'extraction, sont repris dans 50µl d'eau distillée. 10µl de la suspension sont prélevés auxquels on ajoute 20µl de tampon porteur concentré 3 fois (Tris-HCl 180mM pH 6,8, contenant 6% sodium dodecyl sulfate, 30% glycérol, 0.006%  
30 de bleu de bromophénol) en présence d'un agent réducteur le dithiothreitol (DTT) à 10 mM final, puis dénaturés pendant 3 minutes à 100°C.

          Les différents échantillons sont déposés sur un gel d'acrylamide à 6% dans une solution de Tris-HCl 0,4M  
35 pH 8,8 contenant 0.1% de sodium dodecyl sulfate, avec un gel de concentration à 4% dans une solution Tris-HCl 0,4M pH6.8 contenant 0.1% sodium dodecyl sulfate. Sur chaque

gel, trois puits sont réservés pour respectivement les poids moléculaires standards (6H Sigma TM), le collagène I témoin (bovin ou humain, 3µg/puits), et l'échantillon correspondant au plant non transformé. La migration  
5 s'effectue dans un tampon de migration Tris 0.025M-glycine 0.2M contenant 0.1% de sodium dodecyl sulfate sous une tension de 100 volts dans le gel de concentration, augmentée à 200 volts pour le gel de séparation.

*Immunotransfert:*

10 Après migration, les gels sont apposés à une membrane poly(vinylidene difluoride) (Immobilon-P, Millipore TM) préalablement mouillée par du méthanol 100% puis imbibée dans le tampon de transfert CAPS (10mM cyclohexylamino-1-propanesulfonic acid, pH11, méthanol  
15 5%). L'ensemble gel plus membrane est placé dans un appareil de transfert (Biorad TM) et soumis à une tension de 60 volts pendant 16 heures, réfrigéré par un système de circulation d'eau froide.

Afin de vérifier l'efficacité du transfert, les  
20 gels sont colorés en bleu de Coomassie R-250 (solution à 0.2% dans le méthanol 40% et l'acide 10%) pendant 30 minutes. Ils sont ensuite décolorés sélectivement dans une solution de méthanol 15% et d'acide acétique 7,5% et les protéines non transférées ou partiellement transférées  
25 sont repérées par la présence de bandes restant sur le gel. Par ailleurs, les membranes sont colorées au Ponceau S (0,2% dans l'acide acétique 1%) pendant 10 minutes puis décolorées dans l'eau distillée. Les différentes protéines ayant transférées correctement sont repérées, la piste  
30 correspondant aux standards de poids moléculaire excisée, séchée à l'air et conservée entre deux feuilles de papier filtre. Le reste des membranes est complètement décoloré.

Les membranes sont ensuite saturées pendant une heure dans du lait en poudre écrémé 10% puis brièvement  
35 rincées en Tris-HCl 25mM pH 7,4, NaCl 150mM, KCl 2.5mM (TBS). Les membranes sont alors incubées avec les anticorps polyclonaux dirigés contre la triple hélice

majeure du collagène I bovin faits chez le lapin dilués au 1:400 dans le tampon TBS contenant 0.05% Tween 20 (TTBS), 1% albumine sérique bovine, pendant 2 heures à température ambiante. Après incubation avec les anticorps, les membranes sont rincées 6 fois 5 minutes en TTBS, puis incubées avec le conjugué dilué au 1:2000 (anti-IgG de lapin fait chez le porc, couplées à la phosphatase alcaline, Dako TM, Danemark) pendant 1 heure à température ambiante. Les membranes sont à nouveau rincées en TTBS (6 fois 5 minutes) puis révélées à l'aide des réactifs du kit APColor (Biorad TM). Les pistes présentant une ou plusieurs bandes aux poids moléculaires attendus correspondent aux positifs, les numéros des plants correspondants sont répertoriés. Tous les positifs sont soumis à un deuxième criblage (gel + immunotransfert).

#### Résultats:

##### pBIOC706

- La construction pBIOC706 présente 40% de positifs sous la forme d'une bande majoritaire migrant au niveau de la chaîne  $\alpha 1(I)$  témoin. Ce résultat suggère que le C-propeptide (environ 30 Kda) de la chaîne  $\alpha 1(I)$  recombinante a été clivé. Les pistes correspondant au plant témoin non transformé ne présentent aucune bande révélée par l'anticorps anti-collagène I.

- les plants 23, 25, 28, 32, 36, 38, 39, 40, 45 sont identifiés comme positifs, les plants 40 et 45 sont déterminés comme étant les plus productifs par leur présence dans les surnageants et leur abondance dans le culots.

##### pBIOC707

- La construction pBIOC707 présente 27% de positifs sous la forme d'une bande majoritaire migrant au dessus de la chaîne  $\alpha 1(I)$  témoin correspondant à un poids moléculaire d'environ 140KDa. La masse moléculaire attendue pour la chaîne recombinante entière étant d'environ 160KDa, ce résultat suggère que le N-propeptide

(environ 20KDa) ou le C-propeptide (environ 30KDa) de la chaîne  $\alpha 1(I)$  recombinante a été clivé. Les pistes correspondant au plant témoin non transformé ne présentent aucune bande révélée par l'anticorps anti-collagène I.

- 5           - les plants 1, 2, 14, 16, 17, 19 sont identifiés comme positifs, les plants 1, 2 et 14 sont déterminés comme étant les plus productifs par leur présence dans les surnageants et leur abondance dans le culots (fig. 2).

10           **2. Immunofluorescence:**

**Méthode analytique**

Les feuilles des plants sont enrobés dans un milieu d'inclusion à froid (Tissue-Tek, Miles, USA) et sont congelés sur la platine refroidie à  $-30^{\circ}\text{C}$  du  
15 cryostat.

Des lames histologiques traitées par l'aminosilane sont préparées selon le protocole suivant :

- Immersion dans une solution de 3-  
20 aminopropyltriethoxysilane 2% dans l'acétone, 5 secondes  
-Lavage à l'acétone, 2X 1 minutes  
-Lavage à l'eau distillée  
-Séchage pendant une nuit à  $42^{\circ}\text{C}$ .

Des coupes à congélation de  $7\mu\text{m}$  sont réalisées  
25 pour tous les plants à l'aide d'un cryostat et déposées sur les lames traitées. Les lames sont alors traitées comme suit:

- Rinçage des lames dans un tampon phosphate (PBS pour phosphate buffer saline, 137 mM NaCl, 2,7mM KCl,  
30 10mM sodium phosphate, pH7,2).

- Blocage des sites aspécifiques par une incubation des lames dans du PBS contenant 1% d'albumine sérique bovine (BSA)

- Incubation dans les anticorps anti-collagène I  
35 (référence 20121, Institut Pasteur, Lyon) dilués au 1:50 dans du PBS-BSA 1% pendant 2 heures à température



ambiante. Des lames contrôles sont réalisées où les anticorps anti-collagène I sont remplacés par du PBS seul ou par un anticorps polyclonal spécifique du collagène IV.

- Rinçage dans du PBS, 3 fois 10 minutes
- 5       - Incubation dans le conjugué (anti-IgG de lapin fait chez la chèvre couplés à l'isothiocyanate de fluoresceine, Biosys TM, France) dilué au 1:300 dans du PBS-BSA 1% pendant 1 heure à température ambiante. Pendant l'incubation et jusqu'à l'observation les lames sont
- 10 placées à l'obscurité.

- Rinçage au PBS, 3 fois 10 minutes
- Traitement des lames à l'Eriochrome Black T 0.3% dans du PBS pour masquer l'autofluorescence des échantillons, pendant 1 minute;
- 15       - Rinçage intensif au PBS
- Monter les lamelles sur les lames avec une goutte de PBS-glycérol 1:1.

L'observation se fait sur un microscope à epifluorescence Zeiss TM Universal équipé du filtre Zeiss

20 TM exciter BP 430-490.

### Résultats

L'observation des lames nous ont permis les conclusions suivantes :

- les amas formés par les chloroplastes auto-fluorescent en rouge mais ne gênent pas l'observation de
- 25 fluorescence jaune-vert des plants positifs.

- Une auto-fluorescence jaune-vert est également observée sur des structures particulières de la plante et bien délimitées (canaux).

30 Les plants positifs sont cependant facilement identifiables par la présence de petits amas très colorés dans les cellules. On remarque également dans les plants correspondant à la construction pBIOC706 ( $\alpha 1$  + C-pro) un réseau filamenteux diffus qui apparaît localisé dans la

35 partie extracellulaire de la feuille.

- On note une parfaite correspondance entre les plants positifs déterminés par la méthode gel+transfert et par immunofluorescence surtout pour les plus positifs c'est-à-dire 1, 2, 14, 40 et 45.

5                   **Extraction en tampon neutre:**

Afin d'améliorer l'extraction, une extraction en milieu neutre également utilisée pour le collagène tissulaire et le procollagène (molécule de collagène possédant encore les N- et C-propeptides) est réalisée sur  
10 les culots des positifs de la précédente extraction: 1, 2, 14, 40 et 45.

Les résidus issus de l'extraction en milieu acide de chaque plant y compris les plants témoins non transformés sont lavés à l'eau distillée deux fois avant  
15 d'être repris par un tampon Tris-HCl 50 mM pH 7.4 contenant 150 mM NaCl et des inhibiteurs de protéases : 1mM phényl méthyl sulfonide (PMSF), 1 mM N-ethyl maléimide (NEM), 2.5 mM acide tetra-acétique diamine éthylène (EDTA), à raison de 1 gramme de poudre de départ pour 4 ml  
20 de solution. Le mélange est agité à 4°C pendant 60 heures. Après centrifugation de chaque échantillon à 20 000 g (rotor 12148, Sigma TM) pendant 30 minutes à 4°C. Les culots sont lavés 2 fois en eau distillée et stockés à -20°C. Une fraction des surnageants (20 µl) est prélevée  
25 pour réaliser un dosage protéique sur l'extrait total. Le reste des surnageants est stocké à -20°C.

**Méthode analytique**

cf ci-dessus

**Résultats**

30                   **1. Les pistes correspondant aux plants transformés avec la construction pBIOC706:**

On retrouve pour les plants 40 et 45 la bande majeure observée dans les surnageants après extraction en milieu acide migrant au niveau de la chaîne  $\alpha 1$  du  
35 collagène I témoin, mais aussi une autre bande d'intensité égale migrant au environ de 140 KDa. Ce résultat témoigne

de la présence d'une chaîne  $\alpha 1(I)$  recombinante complète correspondant à la construction pBIOC706. En effet la bande supérieure correspond à la masse moléculaire attendue pour la chaîne  $\alpha 1(I)$  recombinante comprenant le C-propeptide.

**2. Les pistes correspondant aux plants transformés avec la construction pBIOC707:**

On retrouve pour chaque plant (1, 2 et 14) la bande majeure observée dans les surnageants après extraction en milieu acide migrant au environ de 140 KDa, mais aussi une autre bande d'intensité égale migrant au environ de 160 KDa. Ce résultat témoigne de la présence d'une chaîne  $\alpha 1(I)$  recombinante complète pour la construction pBIOC707. En effet la bande supérieure correspond à la masse moléculaire attendue pour la chaîne  $\alpha 1(I)$  recombinante comprenant les N- et C-propeptides.

Pour les deux constructions pBIOC706 et pBIOC707, les pistes correspondant aux culots après extraction et centrifugation des échantillons ne présentent plus aucune bande révélée par l'anticorps anti-collagène I, démontrant l'efficacité de l'extraction en pH neutre (fig3).

**EXEMPLE 7 : CONFORMITE DU COLLAGENE**

**Détermination de la présence des propeptides N- et C-terminaux dans les chaînes  $\alpha 1(I)$  recombinantes extraites en milieu acide:**

L'extraction des plants transformés avec la construction pBIOC707 en pH acide ne permet l'extraction que d'une seule bande migrant légèrement au dessus de la bande  $\alpha 1(I)$  témoin. La masse moléculaire attendue pour le polypeptide recombinant entier étant de 160KDa, l'hypothèse d'un clivage lors de la maturation de la chaîne  $\alpha 1(I)$  recombinante est envisagée. Il en est de même

pour la construction pBIOC706 pour laquelle une bande majoritaire est observée au niveau de la bande  $\alpha 1(I)$  témoin. Cette migration ne peut s'expliquer que par un clivage d'une partie de la chaîne recombinante. Afin d'appuyer notre hypothèse selon laquelle les bandes majoritaires obtenues pour les constructions pBIOC706 et pBIOC707 correspondent aux chaînes  $\alpha 1(I)$  recombinantes dont le C-propeptide aurait été clivé au cours de la production des chaînes collagéniques, un gel d'électrophorèse en milieu réducteur ou non sera réalisé sur un plant positif de chaque construction et analysé après immunotransfert. En effet le N-propeptide lorsqu'il est correctement replié comprend des pontage intrachânes entre les cystéines alors que les cystéines du C-propeptide sont pontées interchânes, processus qui initie la formation de la triple hélice.

#### Méthode analytique

fois 500  $\mu$ l des plants 2 (positif de la construction pBIOC707) et 40 (positif de la construction pBIOC706) sont précipités par l'addition de 50  $\mu$ l d'acide trichloroacétique 100% (TCA) pendant 30 minutes à température ambiante. Les échantillons sont centrifugés à 20 000g pendant 15 minutes à 4°C (rotor 12148, Sigma TM), les culots sont lavés 2 fois en éthanol 100% froid. Les précipités sont alors repris par 20  $\mu$ l de tampon porteur (Tris-HCl 60mM pH 6,8, contenant 2% sodium dodecyl sulfate, 10% glycérol, 0.002% de bleu de bromophénol) en alternant une piste en présence de DTT 10 mM final et une piste sans DTT, puis dénaturés pendant 3 minutes à 100°C. Les gels sont analysés après immunotransfert.

#### Résultats

Après transfert et révélation avec les anticorps anti-collagène I (référence 20121, Institut Pasteur, Lyon), on observe que la chaîne recombinante du plant 2 migre en présence d'agent réducteur légèrement plus lentement que la chaîne recombinante en l'absence d'agent

réducteur. Cette faible différence de migration n'est pas observée pour la chaîne recombinante du plant 40 (construit délété en N-propeptide). On en déduit que cette différence de migration est due au correct repliement du N-propeptide permettant les pontages intrachaine des cystéines. Le C-propeptide serait donc absent des deux polypeptides recombinants (fig. 4, partie A)

**Présence d'un pontage interchaîne entre les cystéines du C-propeptide:**

10 Un gel d'électrophorèse en présence ou non d'agent réducteur est réalisé sur les échantillons extraits en pH neutre pour lesquels on obtient deux bandes majoritaires dont la supérieure doit comprendre le C-propeptide. En effet, cette bande supérieure migre à la  
15 masse moléculaire attendue pour le polypeptide non clivé qu'il s'agisse de la construction pBIOC706 (triple hélice majeure + C-propeptide) ou pBIOC707 (N-propeptide + triple hélice majeure + C-propeptide) soit 140 KDa et 160 KDa respectivement. Dans le cas d'un pontage entre les  
20 cystéines de trois chaînes, étape nécessaire à l'initiation de la formation de la triple hélice, la migration de la bande supérieure non réduite devrait migrer au niveau des bandes  $\gamma$  du collagène I témoin correspondant à la migration de trois chaînes non  
25 dissociées (environ 300 KDa).

**Méthode analytique**

Les essais ont été réalisés sur les plants 1, 2 et 14 (construction pBIOC707) et les plants 40 et 45 (construction pBIOC706) extraits en tampon neutre. 2 fois  
30 150  $\mu$ l de chaque échantillon soit 7,5 $\mu$ g à 11 $\mu$ g de protéines totales par prise d'essai, sont prélevés et précipités par 15 $\mu$ l d'acide trichloroacétique 100% (TCA) comme indiqué précédemment. Une série d'échantillon est alors reprise par 20 $\mu$ l de tampon porteur en l'absence  
35 d'agent réducteur, l'autre série par 20 $\mu$ l de tampon porteur en présence de 10mM de DTT. Sur chaque gel deux

puits sont réservés pour respectivement les poids moléculaires standard (6H Sigma TM), le collagène I témoin (bovin ou humain, 3µg/puits). La migration s'effectue dans un tampon de migration Tris 0.025M-glycine 0.2M contenant  
5 0.1% de sodium dodecyl sulfate sous une tension de 100 volts dans le gel de concentration, augmentée à 200 volts pour le gel de séparation.

*Immunotransfert:*

cf ci dessus dans la partie Criblage par analyse  
10 des plants par électrophorèse suivie d'un immunotransfert.

**Résultats**

**1. Les pistes correspondant aux plants transformés avec la construction pBIOC707:**

En présence de l'agent réducteur, on retrouve  
15 pour chaque plant 1, 2 et 14, deux bandes majeures l'une migrant aux environs de 140 KDa, l'autre aux environs de 160 KDa. En absence d'agent réducteur, la bande à 140KDa migre légèrement plus rapidement (pontage intrachaine des cystéines contenues dans le N-pro lui conférant une  
20 structure plus compacte et donc migrant plus vite dans le gel que la forme réduite non-repliée). La bande à 160KDa en l'absence d'agent réducteur disparaît et apparaît une bande majoritaire migrant aux alentours de 300 KDa (légèrement au dessus des bandes  $\gamma$  du collagène I témoin).  
25 Ce résultat témoigne de la présence de pontage interchaîne entre les cystéines du C-propeptide de la chaîne  $\alpha 1(I)$  recombinante (fig. 4 partie B).

**2. Les pistes correspondant aux plants transformés avec la construction pBIOC706:**

30 En présence de l'agent réducteur, on retrouve pour les plants 40 et 45, deux bandes majeures l'une migrant au niveau de la chaîne  $\alpha 1$  du collagène I témoin, l'autre aux environs de 140 KDa. En absence d'agent réducteur, la migration de la bande inférieure n'est pas  
35 modifiée (confirmant l'absence du C-propeptide). La bande

à 140KDa en l'absence d'agent réducteur disparaît partiellement et apparaît une bande majoritaire migrant aux alentours de 300 KDa (au niveau des bandes  $\gamma$  du collagène I témoin). Ce résultat témoigne de la présence  
5 de pontage interchaînes entre les cystéines du C-propeptide de la chaîne  $\alpha 1(I)$  recombinante (fig. 4 partie C).

**EXEMPLE 8: Formation de la triple hélice**  
Digestion protéolytique

10 Une digestion ménagée à la trypsine ou à la pepsine permet d'attester la formation d'une triple hélice et donc de l'homotrimère  $[\alpha 1(I)]_3$ . En effet, les extrémités N- et C- terminales sont sensibles aux protéases alors que la partie centrale de la molécule,  
15 lorsque la triple hélice est formée, est résistante. Les conditions de digestion permettant la digestion du collagène I dénaturé sans altérer le collagène I natif sont retenues. Les échantillons extraits en milieu neutre sont digérés par la trypsine selon un rapport enzyme-  
20 substrat de 1:20, à 25°C pendant 10 minutes. La réaction est arrêtée par addition de TCA 10% final suivie de deux lavages éthanol 100% froid. Les échantillons sont repris dans du tampon échantillon contenant 10mM DTT, chauffés à 100°C pendant 3 minutes et déposés sur un gel  
25 d'électrophorèse 6%. Les gels sont colorés au bleu de Coomassie ou immunotransférés et révélés par les anticorps anti-collagène I (référence 20121, Institut Pasteur Lyon) comme indiqué précédemment. Les plants pour  
30 lesquels la triple hélice a été formée permettront l'observation d'une bande unique migrant au niveau de la chaîne  $\alpha 1(I)$  témoin.

La formation de la triple hélice peut-être également mise en évidence par l'utilisation de la pepsine. Dans ce cas les échantillons sont dialysés contre  
35 de l'acide acétique 0.5M, 200mM NaCl, pH 2.5 puis soumis à

une digestion par la pepsine. Les conditions de digestion utilisées sont celles qui permettent la digestion complète de collagène I témoin dénaturé. La réaction est arrêtée par neutralisation du milieu de digestion. Les  
5 échantillons sont alors précipités au TCA 10% final et traités comme indiqué ci-dessus.

On peut également utiliser les digestions aux protéases pour déterminer la température de dénaturation du collagène recombinant et donc la stabilité de la triple  
10 hélice formée dans les plantes. Pour cela les échantillons sont digérés par la trypsine à des températures de réaction croissantes (de 25°C à 45°C), puis traités comme indiqués ci-dessus. L'apparition de bandes correspondant à de la dégradation protéolytique indique la température de  
15 dénaturation du collagène recombinant. La température du collagène I natif est de 41°C; celle de l'homotrimère  $\alpha 1(I)$  recombinant réalisé dans des cellules eucaryotes est similaire à celui du collagène I témoin mais toutefois, l'homotrimère est partiellement clivé dès une température  
20 de 38°C (Geddis et al., 1993).

Après digestion à la trypsine du collagène extrait des plants transformés avec la construction pBIOC707 et pBIOC706 comme décrit précédemment, on observe la persistance d'une bande migrant au niveau de la chaîne  
25  $\alpha 1(I)$  du collagène I témoin. Ce résultat montre que le collagène homotrimérique extrait des plants transformés est replié en triple hélice (Fig 1, partie A).

#### Ombrage tournant

La formation de la triple hélice peut-être  
30 également attester par observation au microscope électronique à transmission de répliques des molécules obtenues après ombrage tournant. En effet, la molécule de collagène lorsqu'elle est repliée en triple hélice se présente sous une forme caractéristique de bâtonnet de 300  
35 nm de long et 1.4 nm de diamètre. Les chaînes  $\alpha 1(I)$  seules



ne sont pas identifiables par cette technique. Les échantillons à une concentration de 10 à 20µg/ml sont dialysés contre du bicarbonate d'ammonium 200 mM pendant une nuit à 4°C, puis mélangés dans un rapport 1:1 avec du glycérol. 5µl sont déposés sur une feuille de mica fraîchement clivée de 1cm<sup>2</sup>, puis placés dans un évaporateur (Med 10, Balzers TM). Environ 2 nm de platine-carbone sont évaporés sur les échantillons à un angle de 8° sous un vide de 2.10<sup>-6</sup> Torr, suivi d'une évaporation de carbone à 90° pendant quelques secondes. Les répliques sont décollées par pénétration à 45° des micas dans un récipient d'eau distillée filtrée et montées sur des grilles de microscopie électronique (600 mesh, cuivre). L'observation se fait sur un microscope à transmission (CM120, Philipps TM).

La formation d'une triple hélice a été confirmée par des observations en microscopie électronique du collagène purifié à partir de plants positifs après ombrage tournant. Les molécules observées se présentent sous la forme caractéristique de bâtonnets de 300 nm de long (Fig 6, partie B). On note l'absence de domaine globulaire à l'une des 2 extrémités de la triple hélice confirmant le clivage du C-propeptide.

#### EXEMPLE 9 :

##### Ultrastructure des plants transformés:

Un fragment de la feuille de chaque plant transformé positif et d'un témoin non transformé sont fixés dans un mélange fixateur glutaraldéhyde 2%, paraformaldéhyde 0.5% dans un tampon citrate-phosphate 0.1M pH 6.8 pendant 8 heures à température ambiante. Pendant la fixation et ultérieurement si nécessaire, les échantillons sont placés dans une cloche à vide pour les dégazer. Les échantillons sont rincés dans le tampon citrate-phosphate plusieurs bains de 15 minutes puis post-fixés dans une solution d'acide osmique 2% dans le tampon

citrate-phosphate 0.1M pendant 2 heures à température ambiante. Les échantillons sont alors déshydratés par des bains successifs d'éthanol de 30° à 100°. L'éthanol absolu est remplacé par un bain d'oxyde de propylène pur, renouvelé 2 fois. La substitution se fait dans des bains contenant un mélange d'oxyde de propylène pur et de résine Epoxy à raison de 1 volume pour 3 volumes respectivement puis 1:1 et finalement 3:1. L'imprégnation se fait dans la résine Epoxy pure sur une nuit et peut se poursuivre en changeant les bains jusqu'à complète immersion des échantillons. Les échantillons sont alors inclus dans des gélules et mis à polymériser pendant 3 jours à 60°C. Des sections ultrafines sont réalisées, contrastées par l'acétate d'uranyle 7% dans du méthanol pendant 10 minutes puis par le citrate de plomb pendant 5 minutes. Les observations se font sur un microscope à transmission électronique (CM120, Philipps TM).

#### EXEMPLE 10 :

##### Purification du collagène recombinant

Le collagène recombinant est purifié à partir des extraits des plantes en utilisant les propriétés physico-chimiques qui lui sont propres. L'acide fait précipiter grand nombre de protéines et les profils électrophorétiques obtenus à partir d'extraction de plants contrôles en milieu acide montrent que peu de protéines endogènes sont extraites. De plus, la chlorophylle n'est pas présente dans ces extraits. Le collagène étant soluble dans l'acide, les différents extraits sont dialysés contre l'acide acétique 0.5M lorsque ceux-ci n'ont pas été extraits directement en milieu acide. Après centrifugation pour éliminer le matériel insoluble, le collagène sera purifié par précipitation fractionnée par les sels en milieu acide : à 0.7M NaCl en acide acétique 0.5M, puis après centrifugation le surnageant est filtré, précipité à 0.9M NaCl en acide acétique 0.5M, puis à nouveau centrifugé pour obtenir le précipité dans le culot. Les

deux précipités 0.7 et 0.9M, contenant l'échantillon sont repris en acide acétique 0.5M puis sont soumis à une électrophorèse sur acrylamide 6% . La pureté des fractions est estimée après coloration des gels en bleu de Coomassie. Pour les échantillons estimés insuffisamment purs, un deuxième cycle de précipitation par les sels peut être réalisé après dialyse de la fraction contre de l'acide acétique 0.5M pour éliminer les traces de sels résiduels. Dans le cas où une étape de purification supplémentaire serait nécessaire, une colonne échangeuse d'ions est réalisée, l'élution se fait par un gradient de NaCl de 0 à 0.5M, les fractions sont soumises à un gel d'électrophorèse comme indiqué précédemment.

Les résultats montrent qu'une majeure partie des protéines des plantes sont éliminées par une étape de précipitation par NaCl 0.4M. Après précipitation par NaCl 0.7M puis 0.9M comme indiquée dans l'exemple 10, le collagène homotrimérique est observé pur à 80-90% dans le précipité NaCl 0.9M .

Un site de liaison à l'héparine situé sur le collagène hétérotrimérique a été décrit dans la littérature (San Antonio et coll., 1994, J. Cell Biol., 125:1179-1188). De tels sites sont responsables de la fixation de protéoglycannes présents dans la matrice extracellulaire mais aussi de l'adhésion des cellules via des récepteurs de type protéoglycannes membranaires. Ces interactions permettent ainsi d'assurer la cohésion du réseau collagénique. La persistance de ce site sur l'homotrimère sera analysée par fixation de l'homotrimère sur colonne d'affinité à l'héparine dans des conditions physiologiques de pH (7.4) et de force ionique (NaCl 0.15M). Cette méthode peut servir de deuxième étape de purification.

35

**EXEMPLE 11 :**

**Composition en acide aminé et  
microséquençage:**

Les différents produits obtenus après extraction des plants transformés sont hydrolysés sous vide par de l'HCl 6N à 115°C pendant 24 heures dans un système PicoTag (Waters, TM). La composition en acide aminés est  
5 déterminée à l'aide d'un analyseur automatique Beckman TM. Cette analyse permet en particulier de vérifier le taux de proline hydroxylée, l'hydroxylation des prolines garantissant la stabilité de la triple hélice.

Le microséquençage s'effectue selon la méthode  
10 d'Edman avec un appareil permettant le séquençage N-terminal automatique des protéines (Applied Biosystems 473 A protein sequencer TM). Les collagènes ou procollagènes recombinants peuvent être chargés sur l'appareil de 2 façons:

15 - après purification, la protéine se présente sous la forme d'une solution concentrée. La protéine est fixée par liaisons hydrophobes sur une pastille de fibre de verre traitée au polybrène. La pastille de verre est directement introduite dans l'appareil.

20 - après électrophorèse et électrotransfert, les bandes d'intérêt sont repérées après coloration au rouge Ponceau 0.02% dans l'acide acétique 0.1% et décoloration à l'eau distillée filtrée, découpées au scalpel puis introduites dans une cellule adaptée à l'appareil.

25 Selon le principe de la méthode d'Edman, chaque cycle libère un acide aminé sous forme d'un complexe phénylthiohydantoine-acide aminé. Ils sont successivement et automatiquement injectés sur une colonne de chromatographie haute pression (HPLC) en phase inverse et  
30 détectés par leur absorbance UV à 269 nm. Par comparaison entre leur temps de rétention et le spectre des standards, on détermine ainsi cycle après cycle la séquence N-terminale de la protéine analysée.

Le collagène obtenu par précipitation à NaCl  
35 0.9M est soumis à une étape supplémentaire de purification (colonne de chromatographie en phase inverse (C8) en HPLC) avant de procéder à l'analyse en acides aminés. L'analyse

obtenue est conforme à celle prédite à partir de l'ADNc de la chaîne  $\alpha 1(I)$  humaine en particulier pour les résidus caractéristiques du collagène, la Glycine et la Proline:

5		Glycine	Proline (Pro+OHPro)
	$\alpha 1(I)$ prédiction ADNc	32.8%	22.7%
	$\alpha 1(I)$ extraite des plantes	28.4%	20.6%

10

D'autre part, afin de vérifier si le clivage du peptide signal de la plante présent dans les constructions pBIOC707 et pBIOC706 s'est effectué correctement mais également pour s'assurer qu'aucune dégradation  
15 protéolytique ne s'est produite en extrémité N-terminale au cours de l'extraction, chacune des bandes inférieures observées sur gel d'électrophorèse pour le collagène extrait des plants transformés avec la construction pBIOC707 et pBIOC706 a été soumise à un séquençage en N-  
20 terminal par dégradation d'Edmann après électrotransfert.

Les séquences obtenues sont: ELAPQLSY pour le collagène correspondant à la construction pBIOC706 et AQVEGQDE pour la construction pBIOC707. Ces résultats montrent que le clivage du peptide signal s'est  
25 effectivement produit et d'autre part confirment que les extrémités N-terminales (le N-télopeptide la construction pBIOC706 et le N-propeptide pour la construction pBIOC707) sont intactes. Par ailleurs, l'ensemble des résultats (exemples 6, 7, 8 et 11) montrent que le  
30 collagène extrait de plantes à maturité correspond à un collagène homotrimérique dont le C-propeptide est clivé.

**Exemple 12: Accumulation du collagène dans les plantes matures.**

Les jeunes plants transformés qui se sont avérés positifs par immunotransfert ont été amenés jusqu'à maturité. Les feuilles des plantes ont été cueillies et, soit directement congelées puis broyées à froid, soit lyophilisées. Quelle que soit la méthode de stockage des feuilles (congelation ou lyophilisation), le collagène est extrait aisément en milieu acide selon les conditions décrites dans l'exemple 6. Après extraction, on retrouve une seule bande majoritaire pour chacun des extraits obtenus à partir de plantes transformés avec la construction pBIOC707 et pBIOC706. Dans chacun des cas cette bande correspond à l'homotrimère après clivage du C-propeptide (voir exemple 6). D'autre part, il est montré que le collagène peut s'accumuler dans les plantes matures et générer ainsi une quantité importante de collagène par gramme de poids frais. Un jeune plant (plant 45 par exemple) peut fournir un extrait contenant environ 6-8µg/ml de collagène homotrimérique alors qu'arrivée à maturité la même plante permettra l'extraction de 40-50µg/ml soit 0.5mg/g de poids frais. On appelle procollagène la molécule de collagène contenant les N- et C-propeptides, pN-collagène et pC-collagène les molécules qui contiennent l'un ou l'autre des domaines terminaux. Enfin le collagène est la forme moléculaire ne présentant que le domaine en triple hélice et correspond dans les tissus animaux à la forme mature. Par conséquent, les produits extraits sont identifiés comme pN-collagène (construction pBIOC707) et collagène (construction pBIOC706).

30

### **Exemple 13: Fibrillogénèse et test d'adhésion**

Le collagène purifié est dilué à 0.2-0.5 mg/ml dans de l'acide acétique 0.1M puis dialysé contre un tampon phosphate salin (PBS), pH 7.4 à 4°C pendant une nuit. L'échantillon est alors placé dans un bain-marie et la température est augmentée progressivement jusqu'à 30°C

35

et stabilisée à cette température pendant 2 heures. La formation des fibres est observée en microscopie électronique après coloration négative ou positive des échantillons. La cinétique de formation des fibres est  
5 suivie par turbidimétrie à l'aide d'un spectrophotomètre, la formation des fibres modifiant la densité optique (Wood et Keech, 1960, Biochem. J., 75: 588-598).

Le collagène I est une protéine adhésive c'est-à-dire que de nombreux types cellulaires sont capables de  
10 le reconnaître et de se fixer dessus de façon spécifique par l'intermédiaire de récepteurs membranaires. Il est possible de tester les propriétés adhésives d'une protéine à l'aide d'un test colorimétrique réalisé en microplaques (Aumailley et coll. 1989, Exp. Cell Res. 181, 463-474).

15 Le collagène extrait est adsorbé pendant une nuit à 4°C sur des plaques multi-puits, puis mis en contact avec une suspension cellulaire à 37°C pendant 20 à 30 minutes. Les cellules non-adhérentes sont éliminées par lavage, les cellules adhérentes sont fixées au  
20 glutaraldéhyde 1%, puis colorées au cristal violet 0.1%. Après lavage intensif des plaques, le colorant fixé par les cellules adhérentes est solubilisé par du Triton X100. Les plaques sont lues sur un lecteur ELISA à 570nm, la valeur de la densité optique est fonction du nombre de  
25 cellules adhérentes.

## REFERENCES BIBLIOGRAPHIQUES

- Amstrong et al. (Malze Handbook; (1994) M. Freeling, V. Walbot Eds.; pp. 665-671).
- 5        - An et al, Plant Physiol, 81, 301-305 (1986)
- Anderson O.D. et Greene F.C., T.A.G., 77, 689-700 (1989).
- Barta et al., Plant Mol. Biol., 6, 347-357 (1986).
- 10       - Bednareck et al Plant cell 3, 1195-1206 (1991).
- Cornelissen et al, Nature, 321, 531-532 (1986)
- Depigny-This et al., Plant. Mol. Biol., 20, 467-479 (1992).
- 15       - Gaubier et al., Mol. Gen. Genet., 238, 409-418 (1993).
- Gedis et al Matrix , 13, 399-405 (1993)
- Holsters et al, Mol Gen Genet, 163, 181-187 (1978)
- 20       - Horsch R.B et al. Science, 227, 1229-1231 (1985).
- Kay et al., Science, 236, 1299-1302 (1987).
- Li S-W et al. Matrix Biology 14, 593-595 (1994)
- 25       - Mc Elroy et al, Mol Gen Genet, 231, 150-160 (1991)
- Marx, Science, 216, 1305-1307 (1982).
- Matsuoka K Proc. Natl. Acad. Sci. USA, 88, 834-838 (1991).
- 30       - Morikawa et al. Biochemistry 19, 2646-2650 (1980)
- Murakami et al., Plant Mol. Biol., 7, 343-355 (1986).
- Ni et al., Plant J., 7, 661-676 (1995).
- 35       - Parkany M, Molecular biomaterials G.W. Hastings, P. Ducheyne eds, CRC Press, Boca Raton, 111-117 (1984)



- Reina et al., Nucleic Acid Research, 18, 6426 (1990).
- Schroeder M.R. et al, Plant Physiol., 101, 451-458 (1993).
- 5       - Vain P. et al., Plant Cell Tissue and Organ Culture, 18, 143-151 (1989).

## REVENDICATIONS

1. Utilisation d'une séquence nucléotidique recombinante contenant, d'une part, un ADNc codant pour  
5 une ou plusieurs chaînes de collagène de mammifère ou les protéines dérivées, et d'autre part, les éléments permettant à une cellule végétale de produire la ou les chaînes de collagène ou les protéines dérivées, codées par ledit ADNc, notamment un promoteur et un terminateur de  
10 transcription reconnus par la machinerie transcriptionnelle des cellules végétales, pour la transformation de cellules de plantes en vue de l'obtention, à partir de ces cellules, ou de plantes obtenues à partir de ces dernières, de la ou des chaînes  
15 de collagène ou des protéines dérivées

2. Séquence nucléotidique recombinante, caractérisée en ce qu'elle contient d'une part la séquence codante pour une ou plusieurs chaînes de collagène ou les protéines dérivés et d'autre part, les éléments permettant  
20 à une cellule végétale de produire la ou les chaînes de collagène ou les protéines dérivés codées par ladite séquence, notamment un promoteur et un terminateur de transcription reconnus par la machinerie transcriptionnelle des cellules végétales.

25 3. Vecteur, notamment plasmide, contenant une séquence nucléotidique selon la revendication 2, insérée en un site non essentiel pour sa répllication.

30 4. Hôte cellulaire, notamment toute bactérie telle que *Agrobacterium tumefaciens*, transformé par un vecteur selon la revendication 3.

5. Procédé d'obtention d'une ou plusieurs chaînes de collagène ou des polypeptides dérivés, caractérisé en ce qu'il comprend:

35 - la transformation de cellules végétales, notamment à l'aide d'un hôte cellulaire selon la revendication 4, lui-même transformé par un vecteur selon la revendication 3, de manière à intégrer dans le génome

de ces cellules une séquence recombinante selon la revendication 2,

- le cas échéant, l'obtention de plantes transformées à partir des cellules transformées susmentionnées,

- la récupération de la ou des chaînes de collagène ou des polypeptides dérivés recombinants produits dans lesdites cellules ou plantes transformées susmentionnées, notamment par extraction, suivie, le cas échéant, par une purification.

6. Plantes, extraits de plantes ou parties de plantes, notamment feuilles et/ou fruits et/ou semences et/ou cellules de plantes, transformés génétiquement, caractérisés en ce qu'elles contiennent une (ou plusieurs) séquence(s) nucléotidique(s) recombinante(s) selon la revendication 2, intégrée(s) de façon stable dans leur génome, ces plantes étant choisies notamment parmi le colza, le tabac, le maïs, le pois, la tomate, la carotte, le blé, l'orge, la pomme de terre, le soja, le tournesol, la laitue, le riz, la luzerne, et la betterave.

7. Chaîne de collagène ou protéine dérivée caractérisée en ce qu'elle est obtenue selon le procédé de la revendication 5.

8. Collagène ou protéine dérivée caractérisé en ce qu'ils sont obtenus selon le procédé de la revendication 5.

9. Produit, notamment gélatine, caractérisé en ce qu'il est obtenu à partir des chaînes de collagène, du collagène ou de leurs protéines dérivées selon les revendications 7 ou 8.

10. Plantes, extraits de plantes ou parties de plantes, notamment feuilles et/ou fruits et/ou semences et/ou cellules de plantes, transformés génétiquement, caractérisés en ce qu'ils contiennent des chaînes de collagène, du collagène ou les protéines dérivées selon les revendications 7 ou 8, ces plantes étant choisies notamment parmi le colza, le tabac, le maïs, le pois, la

tomate, la carotte, le blé, l'orge, la pomme de terre, le soja, le tournesol, la laitue, le riz, la luzerne, et la betterave.

11. Utilisation de plantes, extraits de plantes  
5 ou parties de plantes selon la revendication 6 ou 10, et/ou de produits selon les revendications 7 à 9, pour l'obtention de compositions pharmaceutiques, médicales, odontologiques, cosmétiques ou biotechnologiques.

12 Biomatériau et composition pharmaceutique,  
10 médicale, odontologique, cosmétique ou biotechnologique. caractérisée en ce qu'ils comprennent des plantes, extraits de plantes ou parties de plantes selon la revendication 6 ou 10 et/ou des produits selon les revendications 7 à 9.

15

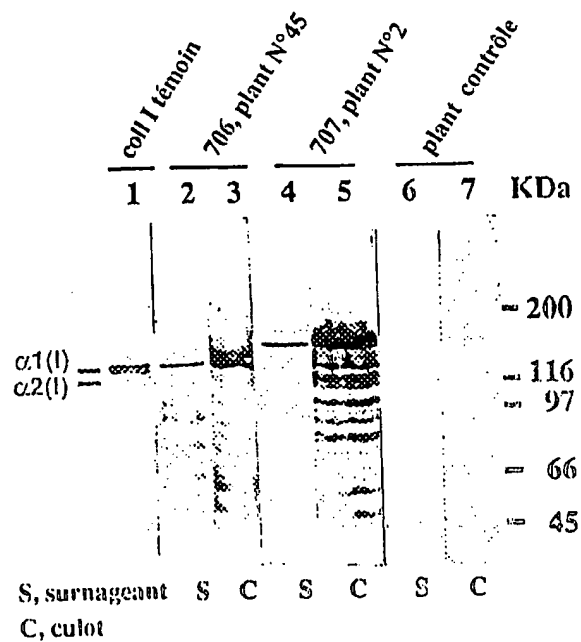


FIGURE 1

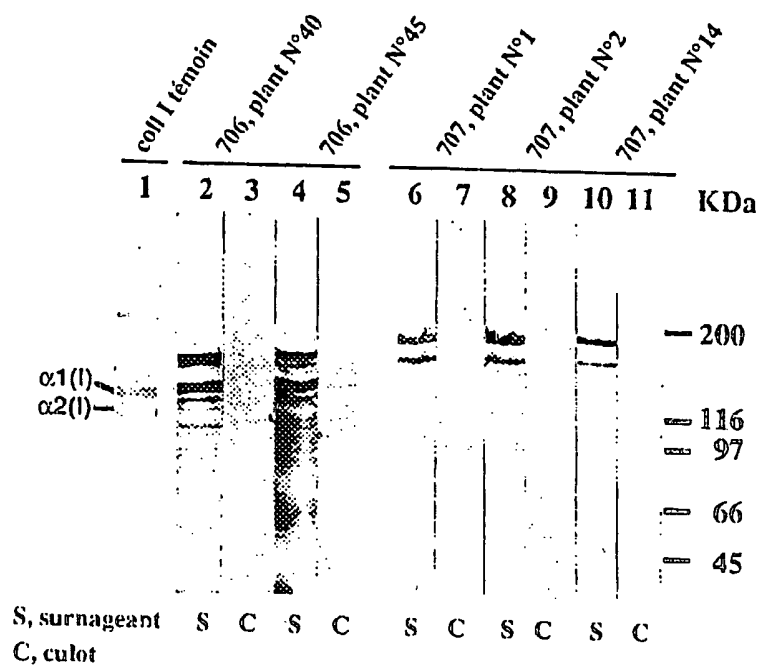


FIGURE 2

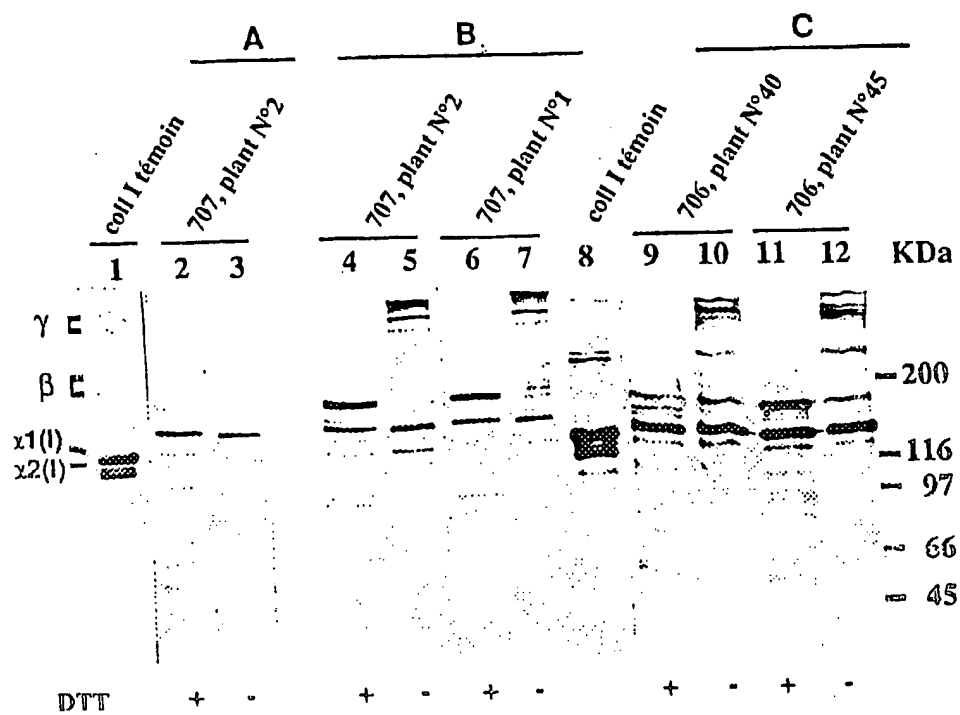


FIGURE 3

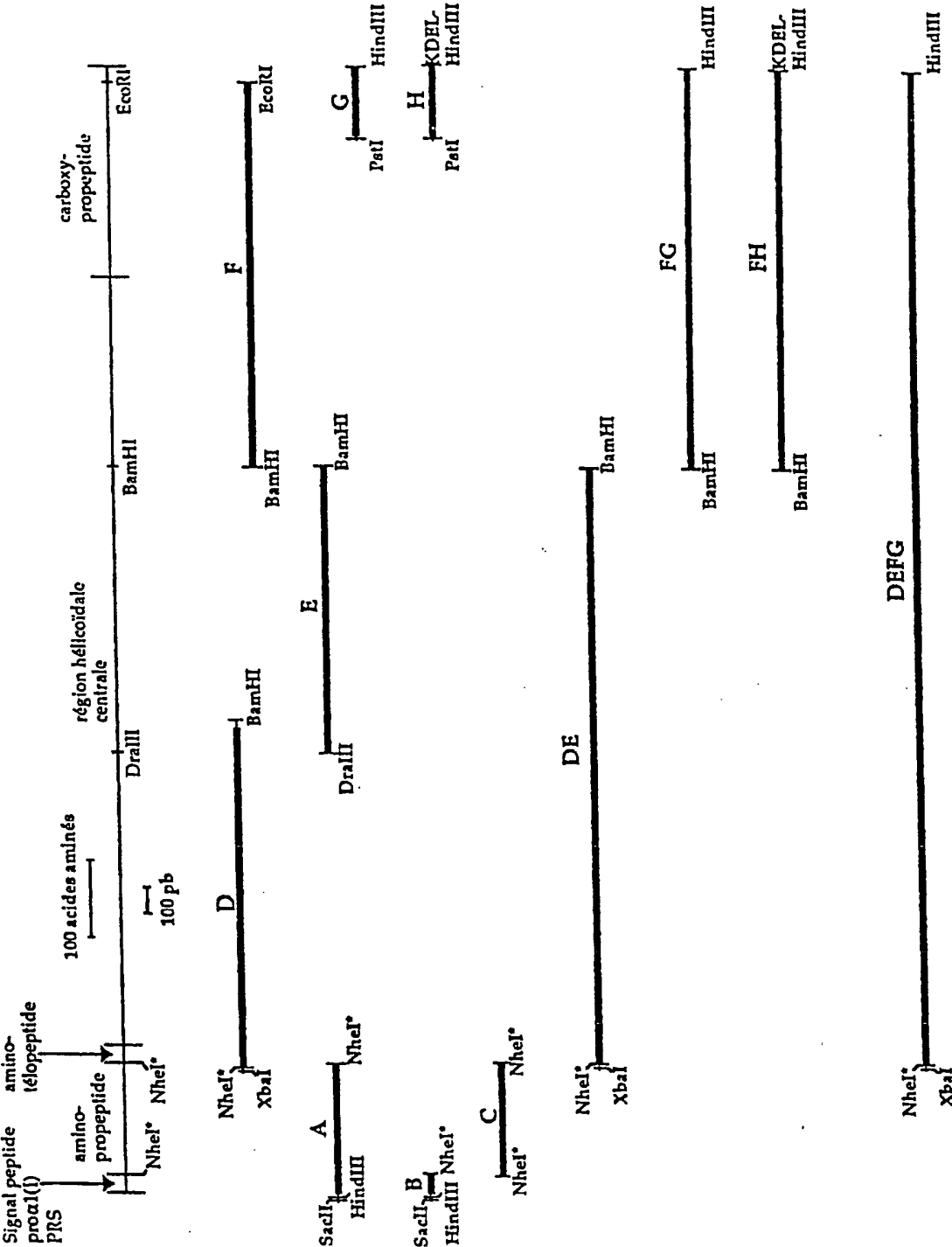


FIGURE 4



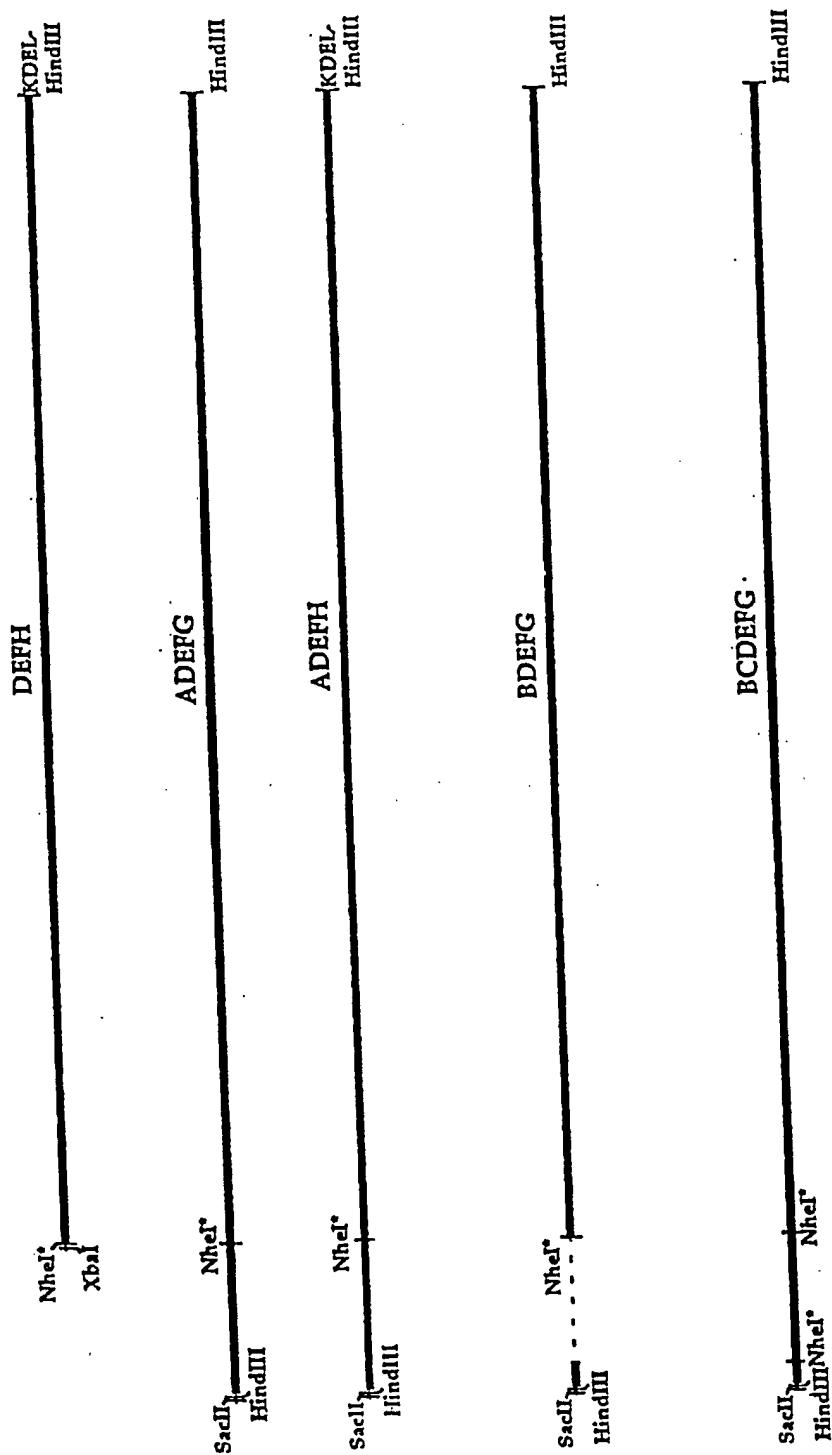


FIGURE 5

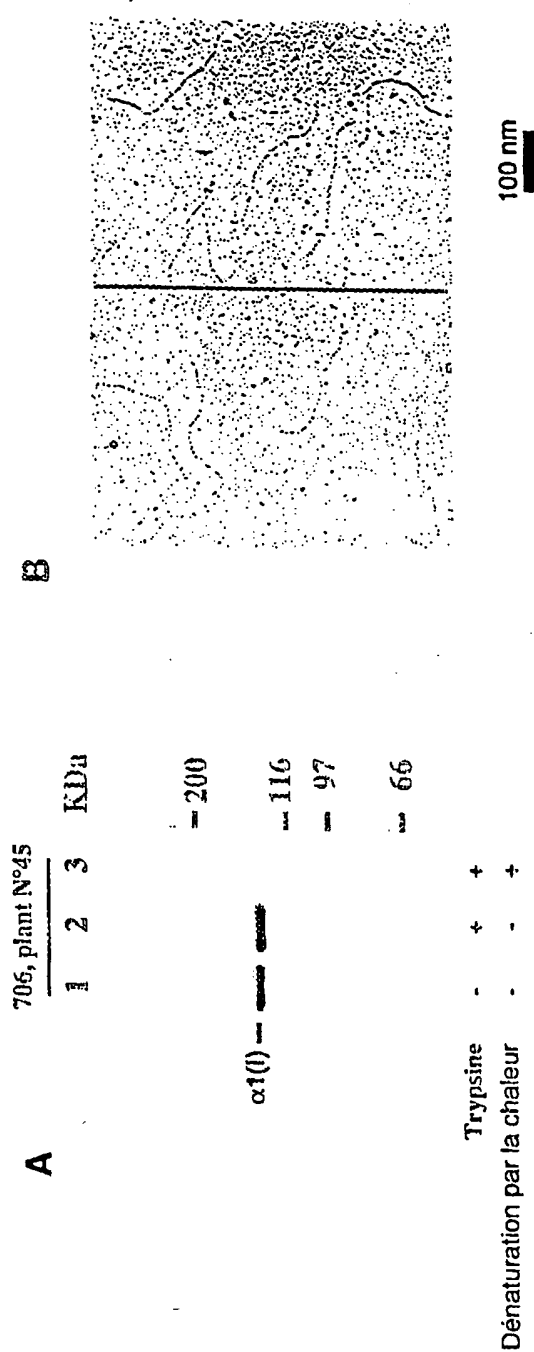


FIGURE 6

Construction ADEFG, proα1(I) complet

Fig.7: a/ séquence protéique

MFSFVDLRLLLLLLAATALLTHGQEEGQVEGQDEDIPPITCVQNGRLRYHDR 50  
 DVWKPEPCRICVCDNGKVLCDVICDETKNCPGAEVPEGECCVCPDGSE 100  
 SPTDQETTGVGEPKGDGTGPRGPRGPAGPPGRDGIPGQPGLPGPPGPPGPP 150  
 GPPGLGGKLAPQLSYGYDEKSTGGISVPGPMGPSGRGLPGPPGAPGPQG 200  
 FQGPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEAGKPGRPGERGPPGPQ 250  
 GARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEFGSPGENGAPGQ 300  
 MGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFFGAVG 350  
 AKGEAGPQGPGRGSEGPQGVRGEPGPPGPAGAAAGPAGNPGADGQPGAKGAN 400  
 GAPGIAGAPGFPGARGPSGPQGPGGPPGPKGNSGEPGAPGSKGDTGAKGE 450  
 PGPVGVQGPFGPAGEEGKRGARGEPGPTGLPGPPGERGGFGSRGFFGADG 500  
 VAGPKGPAGERGSPGPAGPKGSPGEAGRPGEAGLPAGAKGLTGSPGSPGPD 550  
 GKTGPPGPAGQDGRPGPPGPPGARGQAGVMGFPGPKGAAAGEPGKAGERGV 600  
 PGPPGAVGPAGKDGEAGAQQPPGPAGPAGERGEQGPAGSPGFQGLPGPAG 650  
 PPGEAGKPGEQGVPGDLGAPGSPGARGERGFPGERGVQGPFGPAGPRGAN 700  
 GAPGNDGAKGDAGAPGAPGSQAPGLQGMPPERGAAGLPGPKGDRGDAGP 750  
 KGADGSPGKDGVRLTGPIGPPGPAGAPGDKGESGPSGPAGPTGARGAPG 800  
 DRGEPGPPGPAGFAGPPGADGQPGAKGE PGDAGAKGDAGPPGPAGPAGPP 850  
 GPIGNVGAPGAKGARGSGAPPGATGFPGAAGRVGPPGPSGNAGPPGPPGP 900  
 AGKEGGKGRGETGPAGRPGEVGPPGPPGPAGEKGSFGADGPAGAPGTPG 950  
 PQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGKQGPSGASGERGPPGPM 1000  
 GPPGLAGPPGESGREGAPGAEGSPGRDGSPGAKGDRGETGPAGPPGAPGA 1050  
 PGAPGPVGPAGKSGDRGETGPAGPAGPVGPAGARGPAGPQGPGRDKGETG 1100  
 EQGDRGIKGHRGFSGLQGPFGPPGSPGEQGPSGASGPAGPRGPPGSAGAP 1150  
 GKDGLNGLPGPIGPPGPRGRTGDAGPVGPFGPPGPPGPPGPPSAGFDFSF 1200  
 LPQPPQEKAHDGGRYRADDANVVRDRDLEVDTTLSLSQQIENIRSPEG 1250  
 SRKNPARTCRDLKMCHSDWKSGEYWIDPNQGCNLDAIKVFCNMETGETCV 1300  
 YPTQPSVAQKNWYISKNPDKRHWVFGESMTDGFQFEYGGQGSDDPADVAI 1350  
 QLTFRLRMSTEASQNITYHCKNSVAYMDQQTGNLKKALLKGSNEIEIRA 1400  
 EGNSRFTYSVTVDGCTSHTGAWGKTVIEYKTTKTSRLPIIDVAPLDVGAP 1450  
 DQEFGFDVGFVCFL 1464

Souligné: signal peptide de la chaîne proα1(I)

Gras: substitution des acides aminés 158 et 159,

NF par KL

Relief: sites de coupures de l'amino- et de la carboxy-protéinase.

Fig. 7 : b/ Séquence nucléique

AAGCTTAGAC

ATGTTTCAGCTTTGTGGACCTCCGGCTCCTGCTCCTCTTAGCGGCCACCGC 50  
CCTCCTGACGCACGGCCAAGAGGAAGGCCAAGTCGAGGGCCAAGACGAAG 100  
ACATCCCACCAATCACCTGCGTACAGAACGGCCTCAGGTACCATGACCGA 150  
GACGTGTGAAAACCCGAGCCCTGCCGGATCTGCGTCTGCGACAACGGCAA 200  
GGTGTGTGCGATGACGTGATCTGTGACGAGACCAAGAACTGCCCCGGCG 250  
CCGAAGTCCCCGAGGGCGAGTGCTGTCCCGTCTGCCCCGACGGCTCAGAG 300  
TCACCCACCGACCAAGAAACCACCGCGTCGAGGGACCCAAGGGAGACAC 350  
TGGCCCCCGAGGCCCAAGGGGACCCGCAGGCCCCCTGGCCGAGATGGCA 400  
TCCCTGGACAGCCTGGACTTCCCGGACCCCCCGGACCCCCCGGACCTCCC 450  
GGACCCCCTGGCCTCGGAGGAAAGCTAGCTCCCCAGCTGTCTTATGGCTA 500  
TGATGAGAAATCAACCGGAGGAATTTCCTGTCCTGGCCCCATGGGTCCCT 550  
CTGGTCCTCGTGGTCTCCCTGGCCCCCTGGTGCACCTGGTCCCCAAGGC 600  
TTCCAAGGTCCCCCTGGTGAGCCTGGCGAGCCTGGAGCTTCAGGTCCCAT 650  
GGGTCCCCGAGGTCCCCCAGGTCCCCCTGGAAAGAATGGAGATGATGGGG 700  
AAGCTGGAAAACCTGGTCGTCTGGTGAGCGTGGGCCTCCTGGGCCTCAG 750  
GGTGCTCGAGGATTGCCCCGAACAGCTGGCCTCCCTGGAATGAAGGGACA 800  
CAGAGGTTTTCAGTGGTTTGGATGGTGCCAAGGGAGATGCTGGTCCCTGCTG 850  
GTCCTAAGGGTGAGCCTGGCAGCCCTGGTGAAAATGGAGCTCCTGGTTCAG 900  
ATGGGCCCCCGTGGCCTGCCTGGTGAGAGAGGTGCCCCCTGGAGCCCCCTGG 950  
CCCTGCTGGTGCTCGTGGAATGATGGTGCTACTGGTGCTGCCGGGCCCC 1000  
CTGGTCCCACCGGCCCCGCTGGTCCCTCCCTGGCTTCCCTGGTGCTGTTGGT 1050  
GCTAAGGGTGAAGCTGGTCCCCAAGGGCCCCGAGGCTCTGAAGGTCCCCA 1100  
GGGTGTGCGTGGTGAGCCTGGTCCCCCTGGCCCTGCTGGTGCTGCTGGCC 1150  
CTGCTGGAAACCCTGGTGCTGATGGACAGCCTGGTGCTAAAGGTGCCAAT 1200  
GGTGCTCCTGGTATTGCTGGTGCTCCTGGCTTCCCTGGTGCCCCGAGGCC 1250  
CTCTGGACCCCAGGGCCCCGGCGGCCCTCCTGGTCCCAAGGGTAACAGCG 1300  
GTGAACCTGGTGCTCCTGGCAGCAAAGGAGACACTGGTGCTAAGGGAGAG 1350  
CCTGGCCCTGTTGGTGTTCAAGGACCCCCCTGGCCCTGCTGGAGAGGAAG 1400

Fig. 7 : b/ Séquence nucléique - suite

AAAGCGAGGAGCTCGAGGTGAACCCGGACCCACTGGCCTGCCCCGGACCCC 1450  
CTGGCGAGCGTGGTGGACCTGGTAGCCGTGGTTTCCCTGGCGCAGATGGT 1500  
GTTGCTGGTCCCAAGGGTCCCGCTGGTGAACGTGGTTCTCCTGGCCCCGC 1550  
TGGCCCCAAAGGATCTCCTGGTGAAGCTGGTCGTCCCGGTGAAGCTGGTC 1600  
TGCCTGGTGCCAAGGGTCTGACTGGAAGCCCTGGCAGCCCTGGTCCTGAT 1650  
GGCAAACTGGCCCCCCTGGTCCCGCCGGTCAAGATGGTCGCCCCGGACC 1700  
CCCAGGCCCACTGGTGGTCCCCGTGGTCAGGCTGGTGTGATGGGATTCCCTG 1750  
GACCTAAAGGTGCTGCTGGAGAGCCCCGGCAAGGCTGGAGAGCGAGGTGTT 1800  
CCCGGACCCCCCTGGCGCTGTCCGTCTGCTGGCAAAGATGGAGAGGCTGG 1850  
AGCTCAGGGACCCCCCTGGCCCTGCTGGTCCCGCTGGCGAGAGAGGTGAAC 1900  
AAGGCCCTGCTGGCTCCCCCGGATTCCAGGGTCTCCCTGGTCTGCTGGT 1950  
CCTCCAGGTGAAGCAGGCAAACCTGGTGAACAGGGTGTTCCTGGAGACCT 2000  
TGGCGCCCCCTGGCCCCCTCTGGAGCAAGAGGCGAGAGAGGTTCCTGGCG 2050  
AGCGTGGTGTGCAAGGTCCCCCTGGTCTGCTGGACCCCGAGGGGCCAAC 2100  
GGTGCTCCCGGCAACGATGGTGCTAAGGGTGATGCTGGTGCCCCCTGGAGC 2150  
TCCCGGTAGCCAGGGCGCCCCCTGGCCTTCAGGGAATGCCTGGTGAACGTG 2200  
GTGCAGCTGGTCTTCCAGGGCCTAAGGGTGACAGAGGTGATGCTGGTCCC 2250  
AAAGGTGCTGATGGCTCTCCTGGCAAAGATGGCGTCCGTGGTCTGACCGG 2300  
CCCCATTGGTCTCTCCTGGCCCTGCTGGTGCCCCCTGGTGAACAAGGGTGAA 2350  
GTGGTCCCAGCGGCCCTGCTGGTCCCCTGGAGCTCGTGGTGCCCCCGGA 2400  
GACCGTGGTGAGCCTGGTCCCCCGGCCCTGCTGGCTTTGCTGGCCCCCC 2450  
TGGTGCTGACGGCCAACCTGGTGCTAAAGGCGAACCTGGTGATGCTGGTG 2500  
CCAAAGGCGATGCTGGTCCCCCTGGGCCCTGCCGGACCCGCTGGACCCCCCT 2550  
GGCCCCATTGGTAATGTTGGTGCTCCTGGAGCCAAAGGTGCTCGCGGCAG 2600  
CGCTGGTCCCCCTGGTGCTACTGGTTTCCCTGGTGGCTGCTGGCCGAGTCG 2650  
GTCCTCCTGGCCCCCTCTGGAAATGCTGGACCCCTGGCCCTCCTGGTCTCT 2700  
GCTGGCAAAGAAGGCGGCAAAGGTCCCCGTGGTGAGACTGGCCCTGCTGG 2750  
ACGTCTCTGGTGAAGTTGGTCCCCCTGGTCCCCCTGGCCCTGCTGGCGAGA 2800  
AAGGATCCCCCTGGTGCTGATGGTCCTGCTGGTGCTCCTGGTACTCCCGGG 2850  
CCTCAAGGTATTGCTGGACAGCGTGGTGTGGTCCGGCCTGCCTGGTCAGAG 2900  
AGGAGAGAGAGGCTTCCCTGGTCTTCCCTGGCCCCCTCTGGTGAACCTGGCA 2950  
AACAAGGTCCCTCTGGAGCAAGTGGTGAACGTGGTCCCCCGGTCCCATG 3000  
GGCCCCCCTGGATTGGCTGGACCCCTGGTGAATCTGGACGTGAGGGGGC 3050  
TCCTGGTGCCGAAGGTTCCCTGGACGAGACGGTTCTCCTGGCGCCAAGG 3100  
GTGACCGTGGTGAGACCGGCCCGCTGGACCCCTGGTGCTCCTGGTGCT 3150  
CCTGGTGCCCCCTGGCCCCGTGGCCCTGCTGGCAAAGAGTGGTGATCGTGG 3200

Fig. 7 : b/ Séquence nucléique - suite

TGAGACTGGTCCTGCTGGTCCCGCCGGTCCCGTCGGCCCCGCTGGCGCCC 3250  
 GTGGCCCCCGCCGGACCCCAAGGCCCCCGTGGTGACAAGGGTGAGACAGGC 3300  
 GAACAGGGCGACAGAGGCATAAAGGGTCACCGTGGCTTCTCTGGCCTCCA 3350  
 GGGTCCCCCTGGCCCTCCTGGCTCTCCTGGTGAACAAGGTCCCTCTGGAG 3400  
 CCTCTGGTCCTGCTGGTCCCGAGGTCCCCCTGGCTCTGCTGGTGCTCCT 3450  
 GGCAAAGATGGACTCAACGGTCTCCCTGGCCCCATTGGGCCCCCTGGTCC 3500  
 TCGCGGTTCGCACTGGTGATGCTGGTCCTGTTGGTCCCCCGGCCCTCCTG 3550  
 GACCTCCTGGTCCCCCTGGTCCTCCCAGCGCTGGTTTCGACTTCAGCTTC 3600  
 CTGCCCCAGCCACCTCAAGAGAAGGCTCACGATGGTGGCCGCTACTACCG 3650  
 GGCTGATGATGCCAATGTGGTTCGTGACCGTGACCTCGAGGTGGACACCA 3700  
 CCCTCAAGAGCCTGAGCCAGCAGATCGAGAACATCCGGAGCCCAGAGGGA 3750  
 AGCCGCAAGAACCCCGCCCGCACCTGCCGTGACCTCAAGATGTGCCACTC 3800  
 TGA CTGGAAGAGTGGAGAGTACTGGATTGACCCCAACCAAGGCTGCAACC 3850  
 TGGATGCCATCAAAGTCTTCTGCAACATGGAGACTGGTGAGACCTGCGTG 3900  
 TACCCCACTCAGCCCAGTGTGGCCCAAGAAGTGGTACATCAGCAAGAA 3950  
 CCCC AAGGACAAGAGGCATGTCTGGTTCGGCGAGAGCATGACCGATGGAT 4000  
 TCCAGTTCGAGTATGGCGGCCAGGGCTCCGACCCTGCCGATGTGGCCATC 4050  
 CAGCTGACCTTCCTGCGCCTGATGTCCACCGAGGCCTCCCAGAACATCAC 4100  
 CTACCACTGCAAGAACAGCGTGGCCTACATGGACCAGCAGACTGGCAACC 4150  
 TCAAGAAGGCCCTGCTCCTCAAGGGCTCCAACGAGATCGAGATCCGCGCC 4200  
 GAGGGCAACAGCCGCTTCACCTACAGCGTCACTGTGATGGCTGCACGAG 4250  
 TCACACCGGAGCCTGGGGCAAGACAGTGATTGAATACAAAACCACCAAGA 4300  
 CCTCCCGCCTGCCCATCATCGATGTGGCCCCCTTGGACGTTGGTGCCCCA 4350  
 GACCAGGAATTCGGCTTCGACGTTGGCCCTGTCTGCTTCCTGTAA 4395  
GCTT

En gras: Signal peptide de la chaîne pro $\alpha$ 1(I)

codon stop TAA

nucléotides substitués, 474, 475 et 477

(CTT par GCA)

En relief: sites de restriction principaux

BamHI, GGATCC

DraIII, CACCTGGTG

EcoRI, GAATTC

SacI, GAGCTC

Souligné: sites de clonages HindIII, AAGCTT, en 5' et 3' de la séquence.

# INTERNATIONAL SEARCH REPORT

Inter. Appl. No.

PCT/FR 97/02331

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/82 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 03051 A (GENE PHARMING EUROP BV ;COLLAGEN CORP (US); KARATZAS COSTAS N (CA)) 8 February 1996 cited in the application see the whole document ---	7,8
X	FR 2 685 347 A (UNIV PASTEUR) 25 June 1993 see the whole document ---	9
Y	WO 93 07889 A (UNIV JEFFERSON) 29 April 1993 see the whole document ---	1-12
Y	WO 95 14099 A (UNIV CALIFORNIA ;RODRIGUEZ RAYMOND L (US)) 26 May 1995 see page 24, line 23 ---	1-12
	-/--	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

### \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

19 May 1998

Date of mailing of the international search report

27/05/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Maddox, A

# INTERNATIONAL SEARCH REPORT

Inter. Appl. No.

PCT/FR 97/02331

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 96 21029 A (UNIV TECHNOLOGIES INT) 11 July 1996 see page 21, line 24 - line 29 see page 24, line 26 ----	1-12
P,X	WO 97 04123 A (GEL TECH GROUP INC) 6 February 1997 see the whole document ----	1-12
E	WO 98 06861 A (AGRIVAX INC ;WELTER LISA M (US)) 19 February 1998 see examples 2-4 ----	1-12
A	GOLDBERG, I., ET AL.: "Cloning and expression of a collagen-analog-encoding synthetic gene in Escherichia coli" GENE, vol. 80, 1989, pages 305-314, XP002042007 see the whole document ----	1-12
A	WO 93 21320 A (UNIV TECHNOLOGIES INT ;MOLONEY MAURICE M (CA)) 28 October 1993 see example 1 ----	1-12
A	GOMORD V ET AL: "PLANTE ET MEDICAMENT, UN NOUVEAU DEPART" BIOFUTUR, vol. 154, March 1996, pages 26-30, XP002033486 see the whole document -----	1-12



# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 97/02331

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9603051	A	08-02-1996	AU 3204795 A CA 2195970 A EP 0772404 A	22-02-1996 08-02-1996 14-05-1997
FR 2685347	A	25-06-1993	NONE	
WO 9307889	A	29-04-1993	AU 680606 B AU 3055592 A AU 4441497 A CA 2121698 A EP 0625048 A JP 7501939 T US 5405757 A US 5593859 A	07-08-1997 21-05-1993 12-02-1998 29-04-1993 23-11-1994 02-03-1995 11-04-1995 14-01-1997
WO 9514099	A	26-05-1995	US 5693506 A AU 1289295 A CA 2176834 A EP 0788550 A JP 9509565 T	02-12-1997 06-06-1995 26-05-1995 13-08-1997 30-09-1997
WO 9621029	A	11-07-1996	US 5650554 A AU 4295096 A CA 2208751 A ZA 9510999 A	22-07-1997 24-07-1996 11-07-1996 13-07-1996
WO 9704123	A	06-02-1997	NONE	
WO 9806861	A	19-02-1998	NONE	
WO 9321320	A	28-10-1993	AU 1583292 A AU 678154 B AU 3884493 A EP 0636179 A EP 0633940 A FI 944550 A FI 944841 A JP 7507199 T JP 7505525 T NO 943664 A	18-11-1993 22-05-1997 08-11-1993 01-02-1995 18-01-1995 30-11-1994 14-12-1994 10-08-1995 22-06-1995 15-11-1994

# INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/FR 97/02331

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9321320 A		NO 943914 A	09-12-1994
		WO 9320216 A	14-10-1993
		US 5650554 A	22-07-1997
<hr/>			

# RAPPORT DE RECHERCHE INTERNATIONALE

Dem. Internationale No

PCT/FR 97/02331

A. CLASSEMENT DE L'OBJET DE LA DEMANDE  
CIB 6 C12N15/12 C12N15/82 A01H5/00

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

## B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 C12N A01H

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)

## C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	WO 96 03051 A (GENE PHARMING EUROP BV ;COLLAGEN CORP (US); KARATZAS COSTAS N (CA)) 8 février 1996 cité dans la demande voir le document en entier ---	7,8
X	FR 2 685 347 A (UNIV PASTEUR) 25 juin 1993 voir le document en entier ---	9
Y	WO 93 07889 A (UNIV JEFFERSON) 29 avril 1993 voir le document en entier ---	1-12
Y	WO 95 14099 A (UNIV CALIFORNIA ;RODRIGUEZ RAYMOND L (US)) 26 mai 1995 voir page 24, ligne 23 ---	1-12
-/--		

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

\* Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"Z" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

19 mai 1998

Date d'expédition du présent rapport de recherche internationale

27/05/1998

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Maddox, A

# RAPPORT DE RECHERCHE INTERNATIONALE

Dem. Internationale No

PCT/FR 97/02331

## C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	WO 96 21029 A (UNIV TECHNOLOGIES INT) 11 juillet 1996 voir page 21, ligne 24 - ligne 29 voir page 24, ligne 26 ---	1-12
P,X	WO 97 04123 A (GEL TECH GROUP INC) 6 février 1997 voir le document en entier ---	1-12
E	WO 98 06861 A (AGRIVAX INC ;WELTER LISA M (US)) 19 février 1998 voir exemples 2-4 ---	1-12
A	GOLDBERG, I., ET AL.: "Cloning and expression of a collagen-analog-encoding synthetic gene in Escherichia coli" GENE, vol. 80, 1989, pages 305-314, XP002042007 voir le document en entier ---	1-12
A	WO 93 21320 A (UNIV TECHNOLOGIES INT ;MOLONEY MAURICE M (CA)) 28 octobre 1993 voir exemple 1 ---	1-12
A	GOMORD V. ET AL: "PLANTE ET MEDICAMENT, UN NOUVEAU DEPART" BIOFUTUR, vol. 154, mars 1996, pages 26-30, XP002033486 voir le document en entier -----	1-12

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Dem .e internationale No

PCT/FR 97/02331

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9603051	A	08-02-1996	AU 3204795 A CA 2195970 A EP 0772404 A	22-02-1996 08-02-1996 14-05-1997
FR 2685347	A	25-06-1993	AUCUN	
WO 9307889	A	29-04-1993	AU 680606 B AU 3055592 A AU 4441497 A CA 2121698 A EP 0625048 A JP 7501939 T US 5405757 A US 5593859 A	07-08-1997 21-05-1993 12-02-1998 29-04-1993 23-11-1994 02-03-1995 11-04-1995 14-01-1997
WO 9514099	A	26-05-1995	US 5693506 A AU 1289295 A CA 2176834 A EP 0788550 A JP 9509565 T	02-12-1997 06-06-1995 26-05-1995 13-08-1997 30-09-1997
WO 9621029	A	11-07-1996	US 5650554 A AU 4295096 A CA 2208751 A ZA 9510999 A	22-07-1997 24-07-1996 11-07-1996 13-07-1996
WO 9704123	A	06-02-1997	AUCUN	
WO 9806861	A	19-02-1998	AUCUN	
WO 9321320	A	28-10-1993	AU 1583292 A AU 678154 B AU 3884493 A EP 0636179 A EP 0633940 A FI 944550 A FI 944841 A JP 7507199 T JP 7505525 T NO 943664 A	18-11-1993 22-05-1997 08-11-1993 01-02-1995 18-01-1995 30-11-1994 14-12-1994 10-08-1995 22-06-1995 15-11-1994

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Dem. Internationale No

PCT/FR 97/02331

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9321320 A		NO 943914 A	09-12-1994
		WO 9320216 A	14-10-1993
		US 5650554 A	22-07-1997
-----			

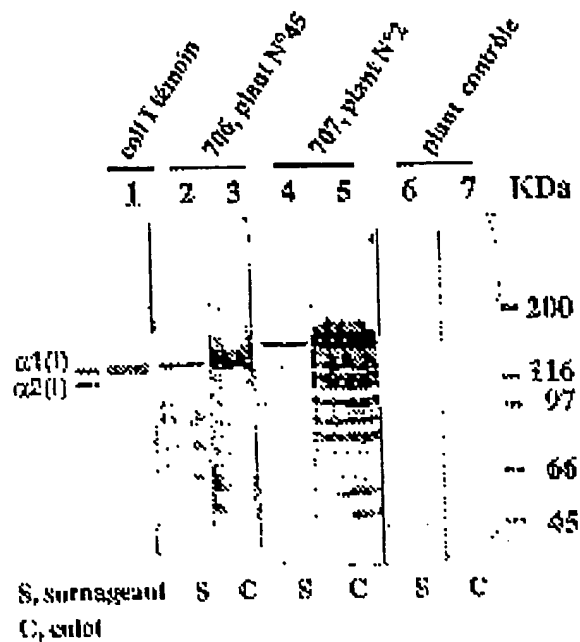


FIGURE 1

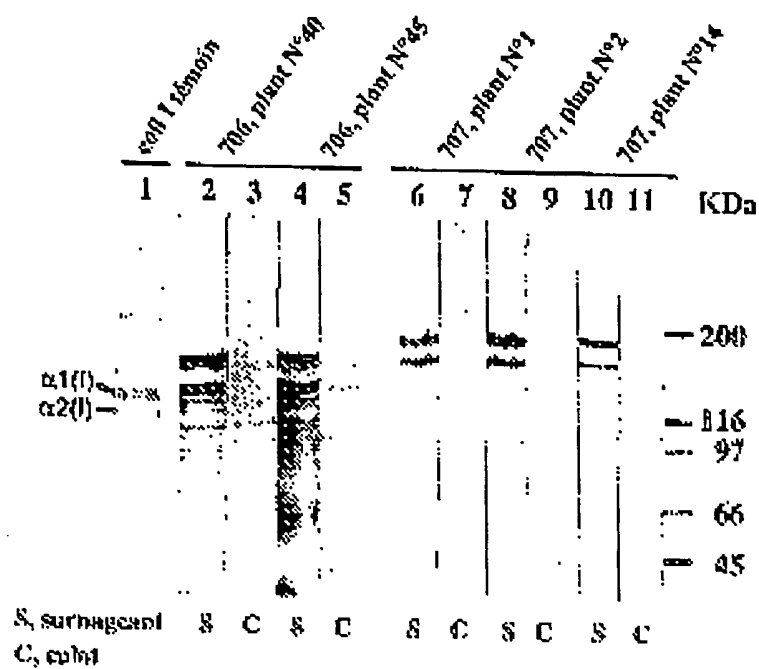


FIGURE 2



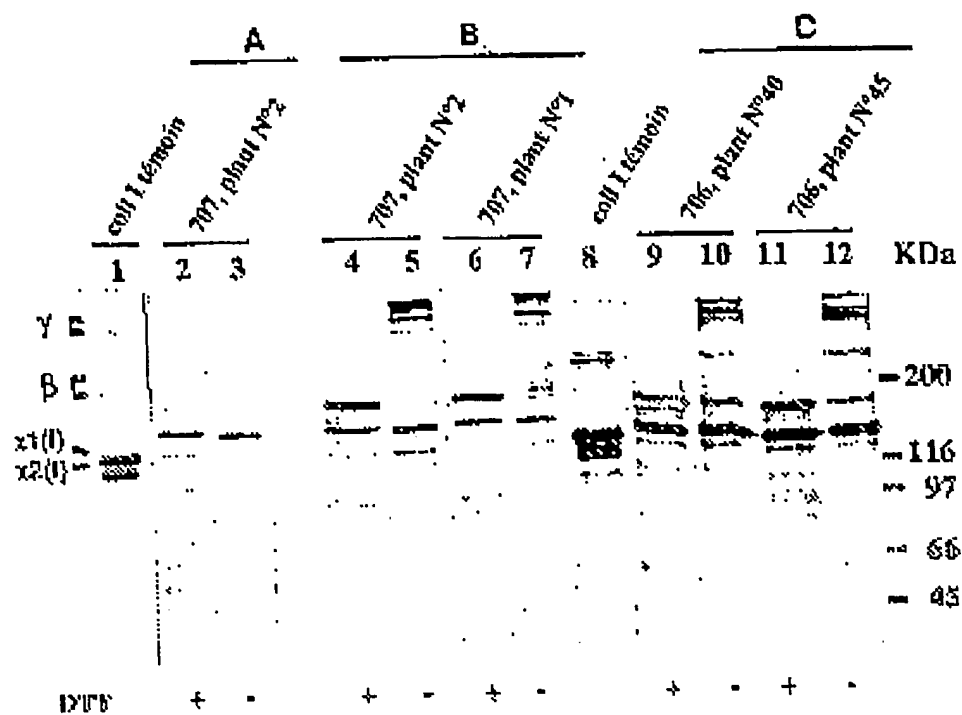
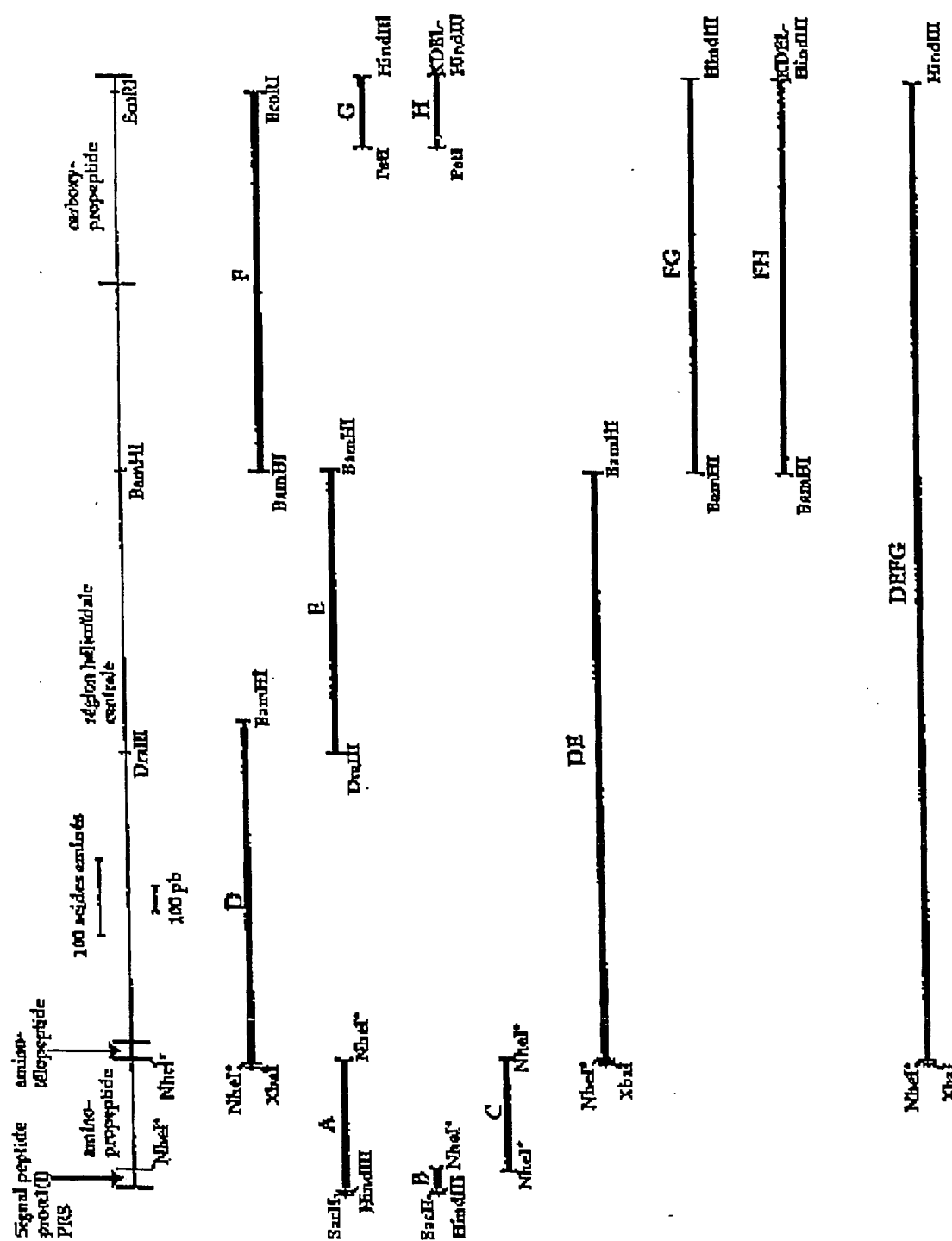


FIGURE 3



**4 EXHIBIT.**

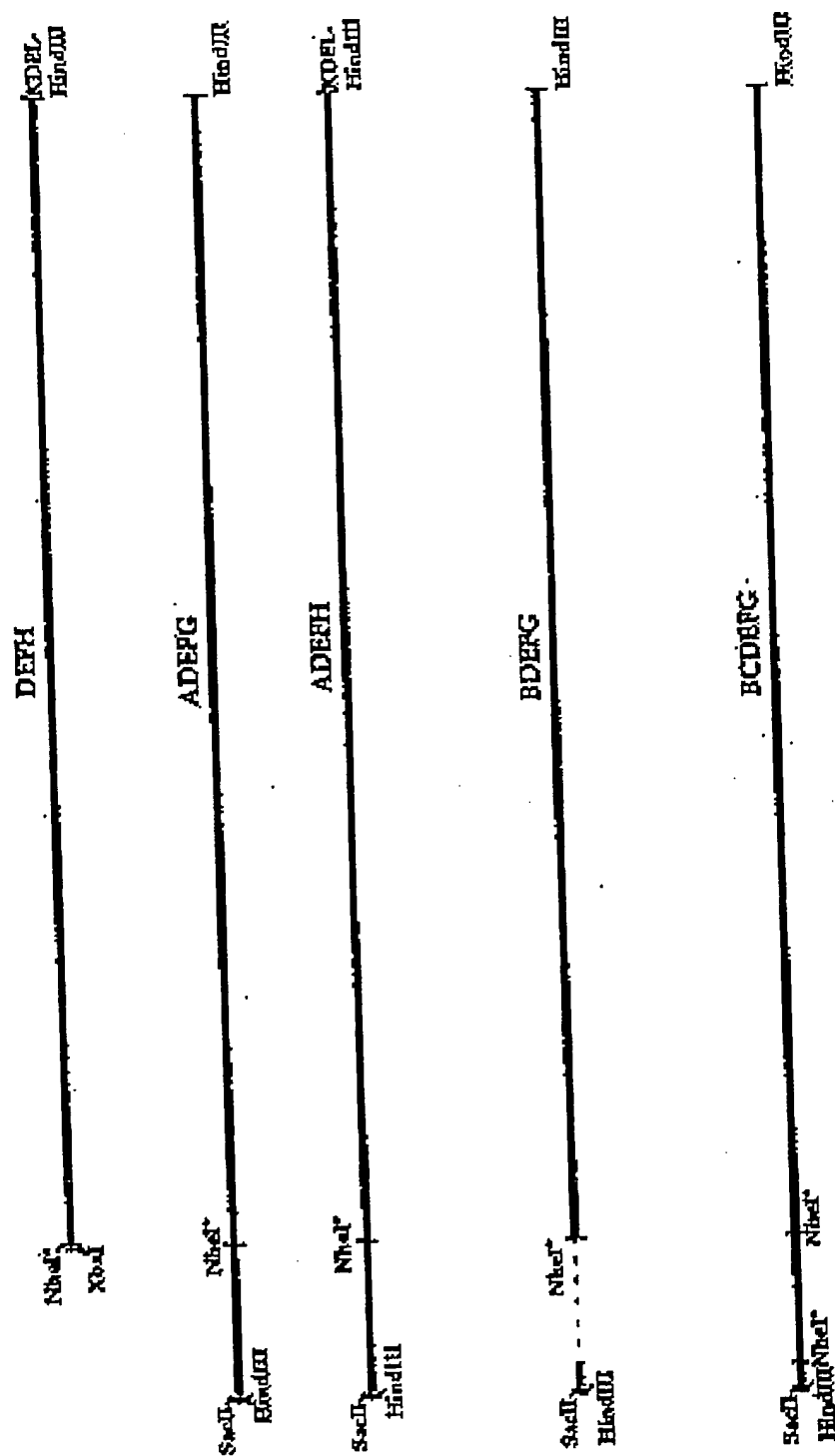


FIGURE 5

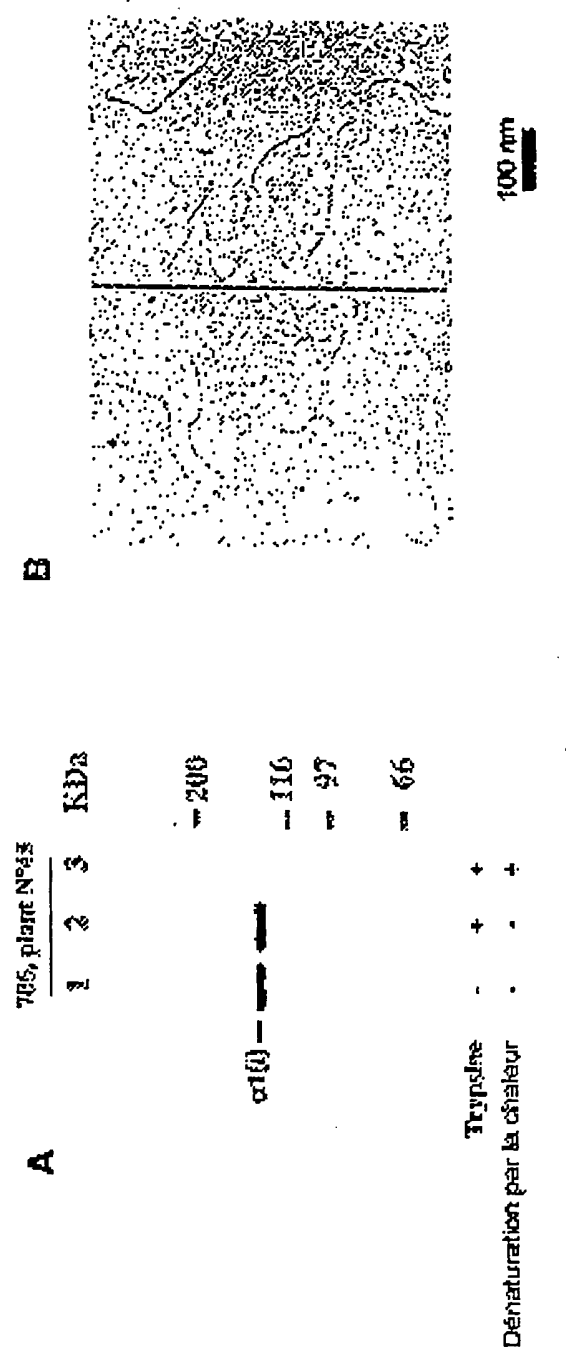


FIGURE 6

Construction ADEFG, pro01(I) complet

Fig.7: a/ séquence protéique

MFSEVDLRLLLLAATALLTHGQEEGQVVGQDEDIPPITCVQNGLYHDR 50  
 DVWKPEPCRICVCDNGKVLCDDEVICDETQNCPGAENVPEGECCPVCFDQSE 100  
 SPTDQETTGVGPKGDTGPRGPRGPAGPFGRDGIPGQPGLPGPFPFPFP 150  
 GPPGLGGKLA**P**QLSYGYDEKSTGGISVFGPMGPSGPRGLEGPFPAGPFQ**G** 200  
 FQGPFGEDGEPGASGPMGPRGFPFPFGKNGDDGEAGKFGRPGERGFP**F**Q 250  
 GARGLPCTAGLPQMKGHRGFSGLDGAKGDAGPAGPKGEFGSPGENGAP**Q** 300  
 MGPRGLPGERGRPGAPGPAGARGNDGATGAAGPRGFTGPAGFPFP**G**AVG 350  
 AKGEAGPQGPFGSEGPGQVRGEPGPPGPAGAAQPA**GN**PGADQPGAKGAN 400  
 GAFQIAGAPGFP**G**ARGPSPGQSPGPPGPKNGSGEPGAPGSKGDTGAKGE 450  
 PGFVG**V**QGPFPAGEEGRGARGEPGPTGLPGFPGERGSPGSRGFPGADG 500  
 VAGPKGPAGERGSFPAGPKGSPGEAGRPQ**E**AGLP**G**AKGLTGSPQSPGPD 550  
 GKTGPPGEPAGQDGRPGPPGPPGARGQAGV**V**GFPGPKGAAGEPGKAGERGV 600  
 PGPPGAVGPAGKDGEAGA**Q**GPPGPAGPAGERGEQGPAGSPGFQGLPGPAG 650  
 PPGEAGKPGEQGVPGDLGAPGPPSGARGERGFPGERGVQGPFPAGPRGAN 700  
 GAFGNDGAKGDAGAPGAPGSQAGPGLQMPGERGAAGLP**G**PKGDRGDAGP 750  
 RGADGSPGKDGVRGLTGPIGPPGPAGAPGDKGESGPPG**P**AGPTGARGAPG 800  
 DRGEFGFPAGFAGFPAGDQPGAKGEFGDAGAKGDAGFP**G**PAGPAGFP 850  
 GPIGNVGAPGAKGARGSGAPPGATGFPGAAGR**V**GFPGPSQ**NA**GPPGPPGP 900  
 AGKEGGKSPRGETQFAGRGFVGFP**GP**PPGAGEKGSFGADGPAGAPGTPG 950  
 PQGIAGQRGVVGLPGQ**R**GERGF**GL**PGPSGEPGKQGP**SG**ASGERGPT**PM** 1000  
 GPPGLAGPPGESGREGAPGAEGSPGRDGSPGAKGDRGETGPAGFP**G**APGA 1050  
 PGAPG**F**VGPAGKSGDRGETGPAGPAGFVG**P**AGARGPAGPQGP**RG**DKGETG 1100  
 EQGDRGIKGHRGFSGLQGPFP**PG**SPGEGQGP**SG**ASGPAGPRGFP**GS**AGAP 1150  
 GKDGLNGLPGPIGPPGFRGRTGDAGFVGPPGPPGPPGPP**PS**AGFDF**SF** 1200  
 LPQFPQEKAHDGGRYVRADANVVRDRDLEVD**IT**LKSL**SQ**QIENIR**S**PEG 1250  
 SEKNPARTCRDLK**CH**SDMKSGEY**W**IDFNQGCN**LA**IKVFCN**ME**TG**ET**CV 1300  
 YPTQPSVAQKNWYISK**NP**KDKRHVWFGE**SM**TDGFQ**F**EYGGQGSDFAD**VA**I 1350  
 QLTFLRLMSTRASON**IT**YHC**NS**VAYMDQQT**EN**LK**AL**LLKG**SN**E**IE**IRA 1400  
 EQNSRFTYSVTVDGCT**SH**TGAWGKT**VI**EYK**TF**KTSRL**FI**IDVAPLDV**G**AP 1450  
 DQETGF**DV**GFV**C**FL 1464

Souligné: signal peptide de la chaîne pro01(I)

Gras: substitution des acides aminés 158 et 159,

NP par KL

Relief: sites de coupures de l'amino- et de la carboxy-protéinase.

Fig. 7 : b/ séquence nucléique

AAGCTTAGAC

ATGTTTCAGCTTTGTGGACCTCCGGCTCCTGCTCCTCTTAGCGGGCCACCGC 50  
 CCTCCTGACGCACGGCCAAAGAGGAAGGCCAAGTCGAGGGGCCAAGACGAAG 100  
 ACATCCCACCAATCACCTGCGTACAGAACGGCCTCAGGTACCATGACCGA 150  
 GACGTGTGGAACCCGAGCCCTGCGGATCTGCGTCTGCGACACCGGCA 200  
 GGTGTGTGCGATGACGTGATCTGTGACGAGACCAAGAACTGCCCGGCG 250  
 CCGAAGTCCCCGAGGGCGAGTGTCTGTCCCTCTGCCCCGACGGCTCAGAG 300  
 TCACCCACCGACCAAGAAACCAACCGGCTCGAGGGACCCAAGGGAGACAC 350  
 TGGCCCCCGAGGCCCAAGGGGACCCCGCAGGCCCCCCCTGGCCGAGATGGCA 400  
 TCCTTGGACAGCCTGGAATTCGCCGACCCCCCGGACCCCCCGGACCTCCC 450  
 GGCACCCCTTGGCCTCGGAGGAAGGCTTGGCTGCCCTGCCCTAGCTGTCTTATGGCTA 500  
 TGATGAGAAATCAACCCGAGGAATTTCCCTGCCCTGGCCCCCATGGGTCCCT 550  
 CTGGTCCCTGCTGCTCTCCCTGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 TTCCAAGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650  
 GGGTCCCGAGGTCCCCAGGTCCCCCTGGAARGAATGGAGATGATGGGG 700  
 AAGCTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750  
 GGTGCTCGAGGATGCTCCCGGACAGCTGGCTCCCTGGAATGAAGGGACA 800  
 CAGAGGTTTTCAGTGTGTTGATGGTCCCAAGGGAGATGCTGGTCTGCTGCTG 850  
 GTCCCTAAGGGTGAGCCTGGCAGCCCTGGTGAAAATGTAAGCTGCTGCTGCTGCT 900  
 ATGGGGCCCCCTGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 950  
 CCTGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000  
 CTGGTCCACCGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050  
 GCTAAGGGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100  
 GGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150  
 CTGCTGGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 GGTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1250  
 CTCTGGACCCCAAGGGCCCGGCGGCCCTCCTGGTCCCAAGGGTAACAGCG 1300  
 GTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1350  
 CCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1400

Fig. 7 : b/ Séquence nucléique - suite

AAAGCGAGGAGCTCGAGGTG/AACCCGACCCACTGACCTGCCCCGGAACCCC 1450  
 CTGGCGAGCGTGGTGGACCTGGTAGCCGTGGTTTCCCTGAGCAGATGGT 1500  
 GTTGCTGCTCCCAAGGGTCCCGCTGCTGAACGTGGTTCCTCGGCCCGC 1550  
 TGCCCCCAAAGGATCTCCTGGTGAAGCTGGTTCCTCCGGTGAAGCTGGTC 1600  
 TGCCTGGTGGCAAGGGTCTGACTGGAAGCCCTGGCAGCCCTGGTCTGAT 1650  
 GGCAAAACCTGGCCCCCTGGTCCCGCCGGTCAAGATGGTGGCCCCGGACC 1700  
 CCCAGGCCCAACCTGGTGGCCCCGTGGTCAAGGCTGGTGTGATGGGATTCCTG 1750  
 GACCTAAAGGTGCTGCTGGAGAGCCCGGCAAGGCTGGAGAGCGAGGTGTT 1800  
 CCGGACCCCCCTGGCGCTGTGGTCTCTGGCAAGATGGAGAGGCTGG 1850  
 AGCTCAGGGAACCCCTGGCCCTGCTGGTCCCGCTGGCGAGAGAGGTGAAC 1900  
 AAGGCCCTGCTGGCTCCCCCGGATTCCAGGGTCTCCCTGGTCTCTGGT 1950  
 CCTCCAGGTGAAGCAGGCAAACTGGTGAACAGGGTGTCTCTGGAGACCT 2000  
 TGCGCCCCCTGGCCCCCTCTGGAGCAAGAGGCGAGAGAGGTTCCTGGCG 2050  
 AGCGTGTGTGCAAGGTCCCCCTGGTCTCTGGACCCCGAGGGGGCCAC 2100  
 GGTGCTCCCGGCAACGATGGTGTGAAGGGTGATGCTGGTGGCCCTGGAGC 2150  
 TCCCGGTAGCCAGGGCGCCCCCTGGCCCTCAGGGGAATGCCCTGGTGAACCTG 2200  
 GTGCAGCTGGTCTTCCAGGGCTTAAGGGTGACAGAGGTGATGCTGGTCCC 2250  
 AAAGGTGCTGATGGCTCTCTCTGGCAAGATGGCGTCCGTGGTCTGACCGG 2300  
 CCCCATTGGTCTCTCTGGCCCTGCTGGTGGCCCTGGTGAACAAGGGTGAAA 2350  
 GTGGTCCAGCGGGCCCTGCTGGTCCCACTGGAGCTCGTGGTGGCCCGGA 2400  
 GACCGTGGTGAGCCCTGGTCCCCCGGCCCCCTGCTGGCTTTGCTGGCCCCC 2450  
 TGGTCTGAOCCGCAACCTGGTGTCTAAAGGCGAACCTGGTGAATGCTGGT 2500  
 CCAAGGCGATGCTGGTCCCCCTGGGCTGGCGGACCCGCTGGAGCCCCCT 2550  
 GGGCCCATTTGGTAATGTTGGTGGTCCCTGGAGGCCAAAGGTGCTGGCGGCA 2600  
 CGCTGGTCCCCCTGGTGGTACTGGTTTCCCTGGTGGTGGTGGCGGATCG 2650  
 GTCTCTCTGGCCCCCTCTGGAAATGCTGGACCCCTGGCCCTCTCTGGTCT 2700  
 GCTGGCAAGAAAGGCGGCAAGGTCCCGCTGGTGAAGACTGGCCCTGCTGG 2750  
 ACGTCTCTGGTGAAGTGGTCCCCCTGGTCCCCCTGGCCCTGCTGGCGAGA 2800  
 AAGGATCCCTGGTGGTGGTCTGGTCTGGTGGTCTCTCTGGTACTCCCGGG 2850  
 CCTCAAGGTATTGCTGGACAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2900  
 AGGAGAGAGAGGCTTCCCTGGTCTTCTCTGGCCCCCTCTGGTGAACCTGGCA 2950  
 AACAAAGGTCCCTCTGGAGCAAGTGGTGAACCTGGTCCCCCGGTCCCATG 3000  
 GGGCCCCCTGGATTGGCTGGACCCCCCTGGTGAATCTGGACGTGAGGGGGC 3050  
 TCTGGTGGCGAAGGTTCCTCTGGACGAGACGGTTCTCTGGGGGGCAAGG 3100  
 GTGAACGTGGTGAAGACCGGCCCCCTGGACCCCTGGTGGTCTCTCTGGTCT 3150  
 CCTGGTGGCCCCCTGGCCCCGTGGCCCTGGTGGCAAGAGTGGTGGTGGTGG 3200

Fig. 7 : b/ Séquence nucléique - suite

```

TGAGACTGGTCCCTGCTGGTCCCCGCCGGTCCCGTGGCCCCCGCTGGCGCCC 3250
GTGGCCCCCGCCGACCCCAAGGCCCCCGTGGTGACAAGGGTGAGACAGGC 3300
GAACAGGSCGACAGAGGCAATMAAGSGTCAACCGTGGCTTCTCTGGCCTCCA 3350
GGGTCCCCCTGGCCCTCCTGGCTCTCTGCTGTAACAAGGTCCCTCTGGAG 3400
CCTCTGGTCCCTGCTGGTCCCCGAGGTCCCCCTGGCTCTGCTGGTGGTCCCT 3450
GGCAAGATGGACTCAACGGTCTCCCTGGCCCCCATTTGGCCCCCTGGTCC 3500
TCGCGGTGCGCACTGGTGATGCTGGTCCCTGTTGGTCCCTCCGGCCCTCCTG 3550
GACCTCCTGGTCCCTGGTCCCTCCAGCGCTGGTTTCGACTTCAGCTTC 3600
CTGCCCCAGCCCTCAAGAGAAGGCTCACGATGGTGGCCGCTACTACCG 3650
GGCTGATGATGCCAATGTGGTTCTGTGACCGTGACCTCGAGGTGGACACCA 3700
CCCTCAAGAGGCTGAGCCAGCAGATCGAGAACATCCGGAGCCCCAGAGGA 3750
AGCCGCAAGAACCCCGCCCGCACCTGCCGTGACCTCAAGATGTGGCACTC 3800
TGACTGGAAGAGTGGAGAGTACTGGATTGACCCCAACCAAGGCTGCAACC 3850
TGGATGCCATCAAAGTCTTCTGCAACATGGAGACTGGTGAGACCTGGGTG 3900
TACCCCACTCAGCCCACTGTGTGGCCAGAACTGGTACATCAGCAAGAA 3950
CCCCAAGGACAAGAGGCATGTCTGGTTCCGGGAGAGCATGACCGATGGAT 4000
TCCAGTTGAGTATGCGCGCCAGGGCTCCGACCCCTGCCGATGTGGCCATC 4050
CAGCTGAACCTTCCCTGCGCCTGATGTCCACCGAGGCCCTCCAGAACATCAC 4100
CTACCACTGCAAGAACAGCGTGGCTTACATGGACCAAGCAGACTGGCAACC 4150
TCAAGAAGGCCCTGCTCCTCAAGGGCTTCAACGAGATCGAGATCCGGGCC 4200
GAGGGCAACAGCCCTTCACTTACAGCGTCACTGTGATGGCTGCCAGAG 4250
TCACACCGGAGCCCTGGGGCAAGACAGTGATTGAATACAAAACCAACAAGA 4300
CCTCCCCGCTGCCCATCATCGATGTGGCCCCCTTGGACGTTGGTCCCCCA 4350
GACCAGCAATTGGCTTCGACGTTGGCCCTGTCTGCTTCTCTTAA 4395
GCTT

```

En gras: Signal peptide de la chaîne proα1(I)

codon stop TAA

nucléotides substitués, 474, 475 et 477

(CTT par GCA)

En relief: sites de restriction principaux

BamHI, GGATCC

DraIII, CACCTGGTG

EcoRI, GAATTC

SacI, GAGCTC

Souligné: sites de clonages HindIII, AAGCTT, en 5' et 3' de la séquence.